820
8841

The Initiation and Propagation of Fatigue Cracks in Mild Steel Pieces of Square Section

H. L. COX, M.A., F.R.Ae.S., A.M.I.Mech.E., and J. E. FIELD
(National Physical Laboratory)

Summary: An investigation has been made to determine the positions and directions of initiation and the directions of propagation of fatigue cracks and to examine the correlation between these positions and directions and the planes on which maximum tensile and maximum shear stresses are generated.

To afford as wide a range as possible of the ratio of maximum shear stress to maximum tensile stress, tests have been made under combinations of alternating bending and torsion; and in order to separate partially the regions of high shear stress from those of high direct stress, the tests have been made on pieces of square section with the plane of bending parallel to one diagonal of the section. Two series of tests have been made; one a preliminary series on pieces having no parallel portion and the other on pieces having a parallel portion about three times the length of the side of the square section. The positions and directions of initiation and the directions of propagation of fatigue cracks have been observed and compared with the positions and directions of the maximum tensile and shear stresses.

Fatigue cracks may be initiated as a result of either high shear stress or high tensile stress and in the present series of tests on mild steel, cracking in tension has occurred in preference to cracking in shear when the ratio of the tensile stress range to the shear stress range has exceeded about 1.6; for values of this ratio less than 1.6, the cracks started in shear (and vice versa); propagation along the plane of maximum shear appears to be preferred up to a slightly greater value of the tensile/shear ratio (about 1.7 possibly). The general direction of a crack formed as a result of high tension usually follows the plane of maximum tension and that of a crack formed as a result of shear usually follows the plane of maximum shear. In detail both types of crack—in this mild steel—deviate quite widely from their general directions but this deviation bears no obvious relation to the microstructure of the material. Cracks propagating along one plane of maximum shear occasionally show a marked tendency to branch along the associated plane of maximum shear; but this tendency is not always observed and in other cases no tendency to branch has been noted.

Paper received May 1952.
[The Aeronautical Quarterly, Volume IV, August 1952]

The present experiments were undertaken to test the feasibility and value of the general method. The results appear encouraging and it is hoped to repeat the tests on other materials and also to elaborate the method a little further.

1. Introduction

In previous investigations into fatigue failure attention has usually been concentrated on the critical conditions of stress magnitude[1], [2] and [3]. Data has been sought on the limiting safe range of stress (or the range of stress corresponding to a pre-selected endurance) and the effect on this limiting range of stress of test-piece form, loading condition and other external circumstance. Certain observations have been made of the direction and manner of initiation and propagation of fatigue cracks[4]; but these observations have usually been incidental to the principal purpose of the tests, and only in work on single crystals has the course of the crack been thoroughly explored[5].

In certain recent tests it has been clear that the fatigue crack may sometimes propagate along planes of principal tension and sometimes along planes of maximum shear, and that a crack initiated in a plane of tension may afterwards propagate along planes of maximum shear. At the same time in many instances the crack has both started and continued along planes of maximum shear. It was desired to pursue these observations and to attempt to establish the conditions under which either of the two types of crack, tensile or shear, might predominate. The N.P.L. Combined Bending and Torsion Fatigue Testing Machine[7] afforded means to impose a wide range of ratios of direct to shear stress; but means was also sought to separate, at least partially, the regions subjected to high shear stress from those under high tensile stress in order to facilitate observation of the start and growth of the two types of crack.

Notation

$K\cos\alpha$	applied range of bending moment
$K\sin\alpha$	applied range of torque
a	length of side of square section
n	fractional distance from one edge of section
f	direct stress due to bending
q	shear stress due to torque
p	function of n, representing distribution of shear stress due to torque
ϕ	$\frac{1}{2}(f^2+4q^2)^{\frac{1}{2}}$ = maximum shear stress
σ	$\frac{1}{2}\{(f^2+4q^2)^{\frac{1}{2}}+f\}$ = maximum direct stress
σ_v	equivalent Mises-Hencky stress
θ	angle of inclination to the axis of the test piece of plane of maximum direct stress
λ	a numerical constant

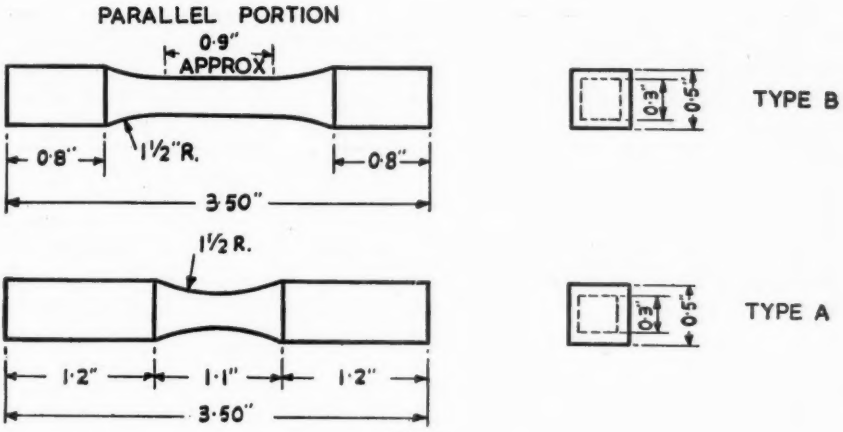

Fig. 1.
Forms of specimen.

2. Choice of Test Piece Form

In the ordinary test piece of circular cylindrical form the shear stress due to torsion is uniform over the whole surface so that, whatever the true criterion of failure, the crack inevitably originates in the region of maximum bending stress. In particular, the regions of maximum tensile and maximum shear stress always coincide and the cause of the crack is evidenced only by its direction and not by its position.

For this reason the present tests have been made on pieces of 0.3 in. square section, subjected to bending in the diagonal plane through two opposite corners of the section. In this form of test piece the direct stress due to bending varies linearly from zero at one edge of each face to a maximum (or minimum) at the other edge, whereas the shear stress due to torsion is zero at each edge and rises to a maximum value in the middle of each face. The region of maximum shear stress varies from the middle of the face under pure torsion, to the outer edge under pure bending, whereas the region of maximum direct stress, which lies also at the outer edge under pure bending, does not move from that site until the applied torsion exceeds the applied bending and then follows the region of maximum shear stress towards the centre of the face, but overtakes it only gradually as the condition of pure torsion is approached. The distributions over the face of the test piece of the maximum shear stress ϕ and the maximum direct stress σ, and the variation of the ratio σ/ϕ are shown in Fig. 2 for various values of the angle α which determines the ratio $\cot \alpha$ of bending to torsion; the gradual shift of the two regions of maximum stresses towards the centre of the face as the condition of pure torsion is approached will be apparent.

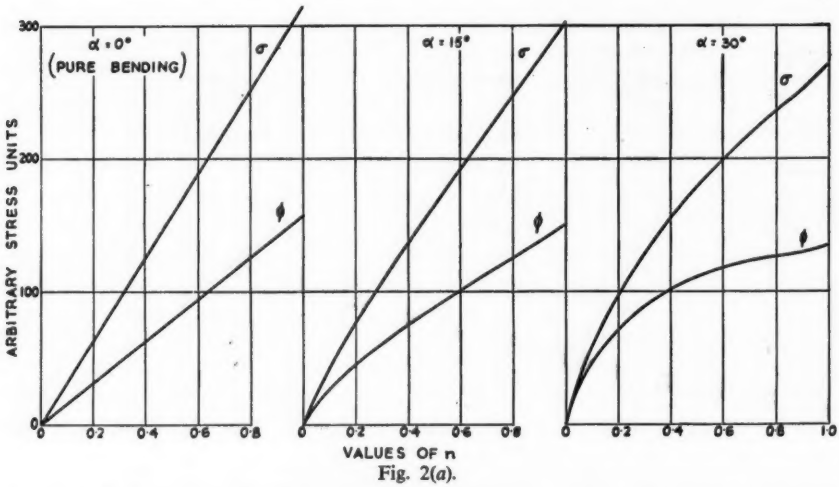

Fig. 2(*a*).

Stress distributions across face of a square section specimen under various combinations of bending and torsion.

(Neutral plane at $n=0$, top or bottom edge at $n=1$)

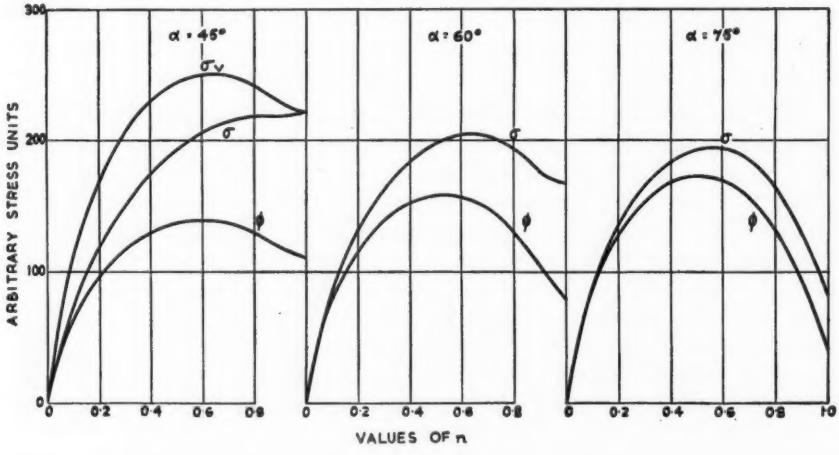

Fig. 2(*b*).

Stress distributions across face of a square section specimen under various combinations of bending and torsion.

(Neutral plane at $n=0$, top or bottom edge at $n=1$)

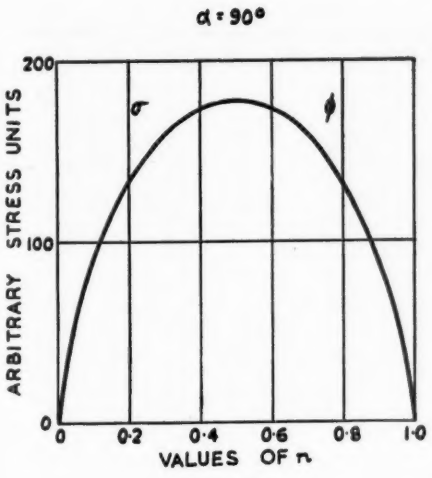

Fig. 2(*c*).

Stress distribution across face of a square section specimen in torsion.

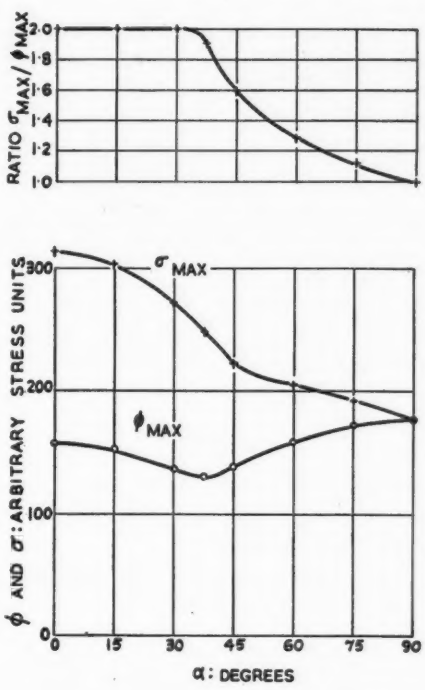

Fig. 2(*d*).

Maximum values of σ and ϕ, and ratio σ/ϕ, as functions of α.

The distributions of stress have been computed by means of the Bernoulli-Euler theory of bending and the St. Venant solution of the torsion problem for a square section[6]. The application of these formulae to the test piece of Type A (Fig. 1), having no parallel portion, is open to some objection, and the course of the crack is certainly affected as it propagates into the enlarged ends of the test piece; but the formulae should be applicable without serious error within the greater part of the parallel length of the pieces of Type B (Fig. 1), and the course of the crack in this central region is unlikely to be affected by the constraint due to the enlarged ends.

3. Preparation of Test Pieces

All the test pieces were cut from mild steel bar drawn from store. Some preliminary tests were made to determine whether the propagation of the fatigue crack was affected by the direction of the finishing treatment applied to the surface of the test piece. These tests showed that so long as the finish was not too coarse the direction of the crack was not influenced. Accordingly all subsequent pieces were finished longitudinally. Pieces A9, A5 and A6 were finished with No. 0 blue black, pieces A2, B2, B4, B5, B6, B7 and B10 with No. 00 blue black, and pieces A7, A8, A10 and B12 were finished with No. 00 blue black and then polished with

an abrasive powder to remove most of the finishing scratches; pieces A2 and A7 were finished transversely, although some longitudinal scratches remained on A7.

4. Testing Procedure

All the tests were done in an N.P.L. Combined Stress Fatigue Testing Machine, Type II, described in Ref. 7. In this machine, although the centre of oscillation of the moving parts is at the centre of the test piece, the bending moment is due to a single load set up at the point of attachment of the loading arm to the rotating disc. Therefore, in the Type B test piece particularly, the stress due to bending varied linearly along the length of the parallel portion, and the actual maximum bending stress was set up slightly beyond one end of this portion in the transition region to the enlarged end. However, this variation of stress along the length of the parallel portion was small, and it appeared to have no appreciable influence on the results except that the cracks tended to start near the fixed end of the test piece and to propagate partially into the enlarged section at that end.

If K is the moment applied in the plane making an angle α with the axis of the test piece, the direct stress due to bending at a point distant na from the neutral plane, measured along the side a of the section ($0 < n < 1$), is $f = 6\sqrt{2}\,(n/a^3)\,K\cos\alpha$, and the shear stress due to torsion is $q = 4.8\,(p/a^3)\,K\sin\alpha$, where p is a function of n defined by St. Venant's solution; the relation between p and n is shown graphically by the curve of ϕ for $\alpha = 90°$ in Fig. 2. The maximum shear stress $\phi = \frac{1}{2}(f^2+4q^2)^{\frac{1}{2}}$ and the maximum direct stress $\sigma = \frac{1}{2}\{(f^2+4q^2)^{\frac{1}{2}}+f\}$; the ratio of ϕ to σ at any point lies always between $\frac{1}{2}$ and 1, but it varies both with α and, for any one value of α, with n also. The plane on which the maximum tensile stress is developed is inclined to the plane perpendicular to the axis of the test piece at an angle θ, where $\tan 2\theta = 2q/f$, and the pair of planes on which the maximum shear stress is developed are inclined at $\pm 45°$ to the plane of maximum tensile stress.

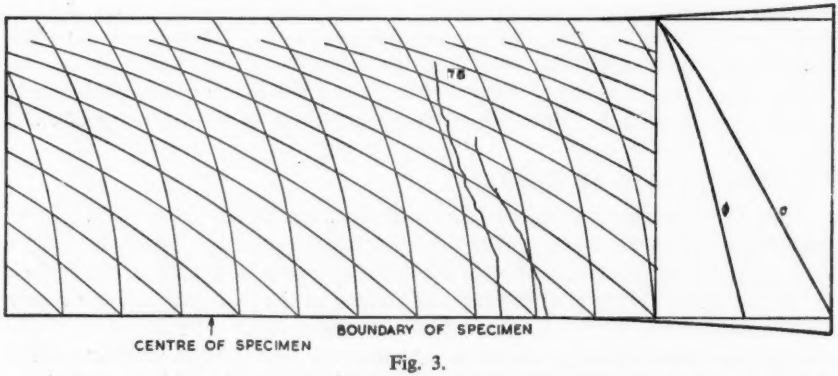

Fig. 3.

Distributions and directions of maximum direct stress and maximum shear stress and courses of fatigue cracks, for $\alpha = 15°$, $f/q = 3.732$. (Showing crack on each of two faces of one specimen.)

FATIGUE CRACKS IN TYPE A SPECIMENS (AFTER MAGNETIC CRACK DETECTION).

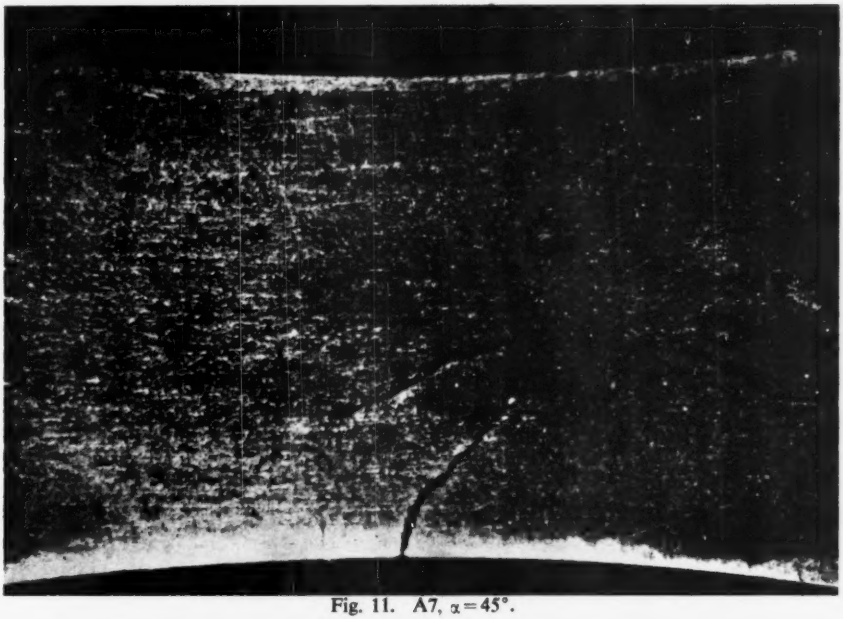

Fig. 11. A7, $\alpha = 45°$.

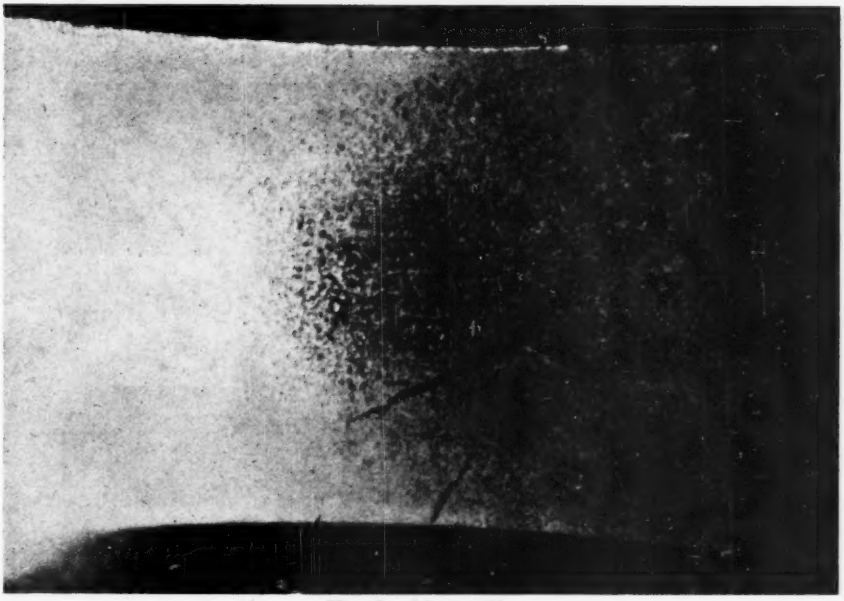

Fig. 12. A8, $\alpha = 45°$.

FATIGUE CRACKS IN TYPE B SPECIMENS (AFTER MAGNETIC CRACK DETECTION).

Fig. 13. B7, $\alpha = 15°$.

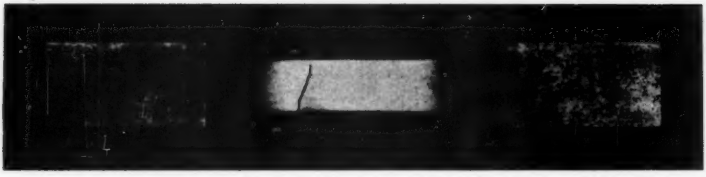

Fig. 14.

Another face of B7, $\alpha = 15°$. Cracks on the other two faces not shown, as they run towards the radius.

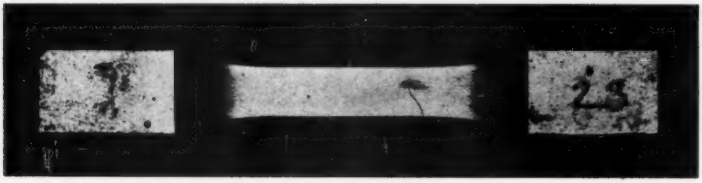

Fig. 15. B2, $\alpha = 30°$.

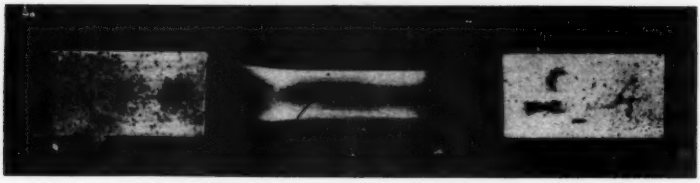

Fig. 16. B8, $\alpha = 45°$.

Library of
Wayne University
Detroit, Mich.

FATIGUE CRACKS IN TYPE B SPECIMENS (AFTER MAGNETIC CRACK DETECTION).

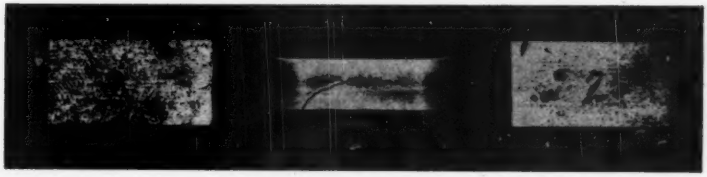

Fig. 17. B9, $\alpha = 45°$.

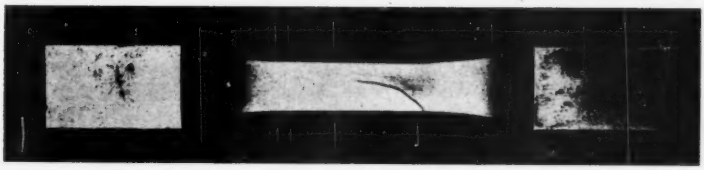

Fig. 18. B6, $\alpha = 45°$.

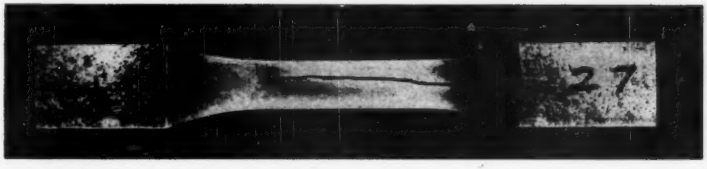

Fig. 19. B4, $\alpha = 60°$.
Rather coarse longitudinal finish may have influenced course of crack.

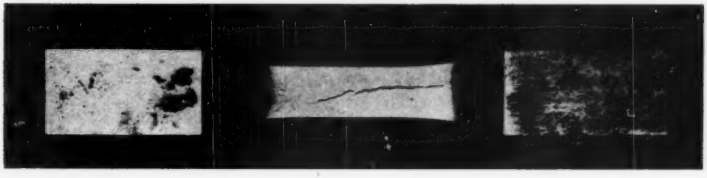

Fig. 20. B10, $\alpha = 60°$.

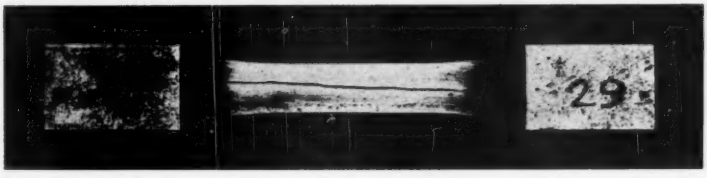

Fig. 21. B5, $\alpha = 75°$.

FATIGUE CRACKS IN TYPE B SPECIMENS. CRACKS ON THE ADJACENT FACE NOT SHOWN, AS THEY RUN TOWARDS THE RADIUS.

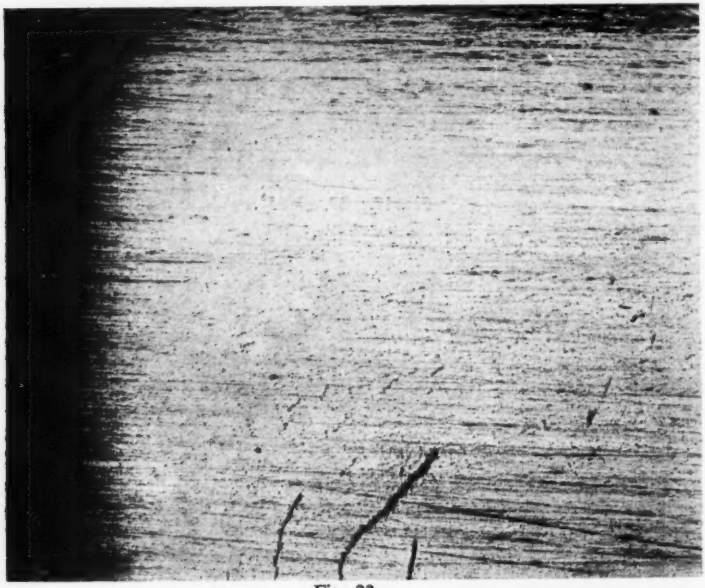

Fig. 22.

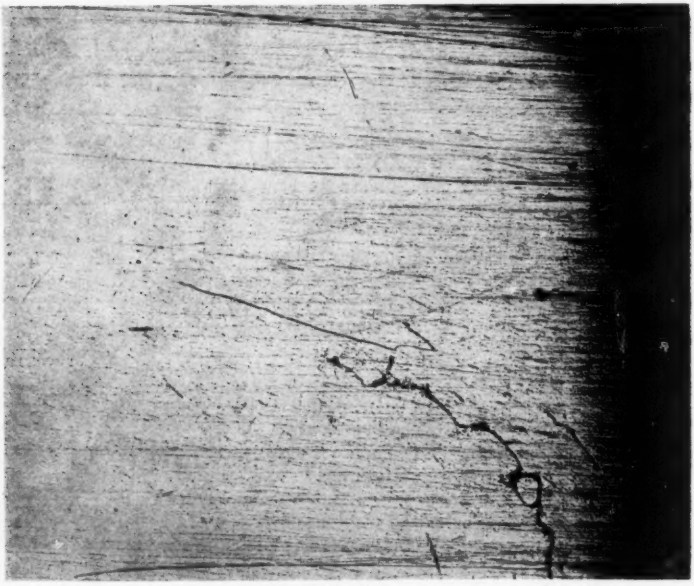

Fig. 23.

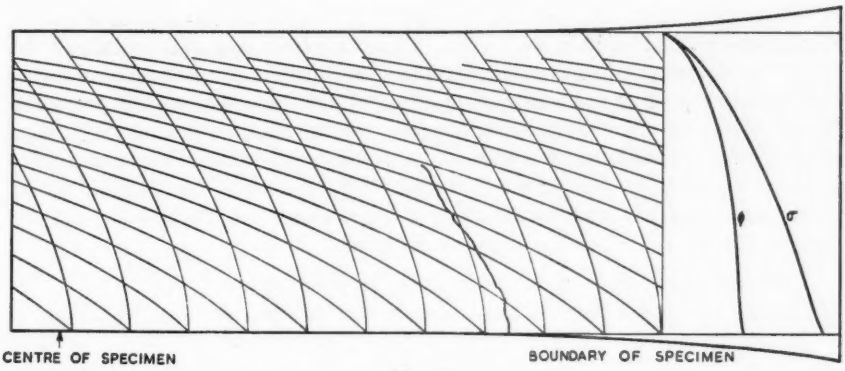

Fig. 4.

Distributions and directions of maximum direct stress and maximum shear stress and course of fatigue crack, for $\alpha = 30°$, $f/q = 1.732$.

Tests were made with values of α equal to 0°, 15°, 30°, 45°, 60° and 90°, one or two pieces being tested at most settings and several at the 45° setting, where the balance between tensile and shear cracking was very close. Owing to variation of mains frequency the speed control of the N.P.L. Combined Stress Machine, which is driven by a synchronous motor, was rather poor. For this reason no attempt was made to maintain stress amplitudes between really close limits; most of the tests were made under fairly high stress ranges so that cracking began at endurances of the order of several hundred thousand to two million reversals and no particular attention was paid to the correlation between the endurance of the test piece and the nominal stress range imposed. These details are summarised in Table I, but the stress ranges there quoted should not be regarded as precise, because speed variation in sympathy with mains frequency may have varied the true values by ± 10 per cent.

5. Observations on Cracks and Correlation with Stress Distributions

Under alternating stresses all four faces of each test piece were subjected to identically the same conditions of loading. Both the maximum direct and shear stresses were developed always in the region between the middle of each face and the upper or lower edge, and both tensile and shear stresses fell away to zero at the edge in the neutral diagonal planes. These distributions of stress, shown in Fig. 2, are repeated at the ends of Figs. 3 to 10. The left hand parts of these figures represent the whole surface of one face of the test piece and the faint lines drawn over the whole of each diagram represent at each point the directions of the plane of maximum direct stress and of one of the pair of planes of maximum shear stress; the direction of the other plane of maximum shear is at right angles.

The darker, more irregular, lines in Figs. 3 to 10 represent the courses of fatigue cracks observed on individual test pieces. These have been projected optically on the basic "co-ordinate" map from photographic plate negatives, the enlargement having been adjusted so that the edges of the face of the test piece in the projection coincided with the edges of the basic diagram. All the cracks which ran to the edge of one face continued across the adjacent face. However, owing to the tendency for these cracks to occur close to the fixed end of the test piece, due to direct stress, these extensions of the crack on the adjacent face soon ran into the enlarged end and therefore their courses are not worth examination.

Photographs of some of the cracked test pieces are shown in Figs. 11 to 23 (*see* pp. 7-10). Except for the very fine cracks in the polished specimen B12, Figs. 22 and 23, these cracks were photographed after magnetic crack detection, and there may exist very small branches which are not shown on the photographs or diagrams.

It will be seen in Figs. 3 and 4 that for values of α not greater than 30°, the fatigue cracks are formed as a result of the range of tensile stress and that they propagate along the planes on which the tensile stress is a maximum; no tendency to deviate into the directions of maximum shear is observed, at least down to ratios of σ to ϕ of 1.6. Similarly, for values of α not less than 60°, Figs. 9 and 10 show that the fatigue cracks result from shear and that they propagate along the plane of maximum shear without any tendency to deviate into the direction of maximum direct stress at least up to ratios of σ to ϕ of 1.3.

With these cracks in shear it is remarkable that the cracks always developed along the nearly axial plane of maximum shear and never along the associated plane at right angles. The only possible explanation for this marked preference for the axial direction seems to be that along the length of the test piece the shear stress is

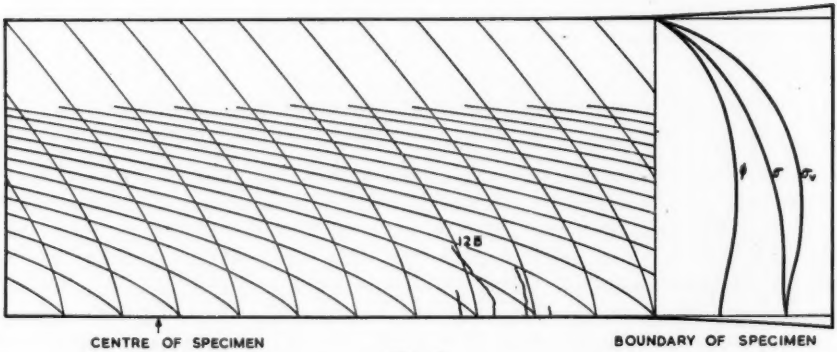

Fig. 5.

Distributions and directions of maximum direct stress and maximum shear stress and courses of fatigue cracks, for $\alpha = 45°$, $f/q = 1$. (Showing cracks on one face of one specimen.)

TABLE I

RESULTS OF COMBINED STRESS FATIGUE TESTS ON MILD STEEL SPECIMENS OF SQUARE SECTION

Specimen	α, *degrees*	$f/q = \cot\alpha$	*Semi-range of applied alternating stress, tons/in.² (maximum values)*				*Endurance* reversals* $\times 10^6$
			Bending stress† component f	*Torsion stress‡ component* q	*Maximum shear stress* ϕ	*Maximum principal stress* σ	
A2	90	0	0	18.0	18.0	18.0	0.22B
A4(1)	0	∞	24.5	0	12.3	24.5	1.93U
(2)	0	∞	29.6	0	14.8	29.6	0.75B
A5(1)	45	1	20.7	11.7	13.0	20.7	1.75U
(2)	45	1	21.8	12.3	13.7	21.8	6.23U
(3)	45	1	23.2	13.1	14.5	23.2	2.01B
A6	45	1	26.0	14.7	16.3	26.0	1.52B
A7	45	1	26.1	14.8	16.4	26.1	0.59B
A8	45	1	25.2	14.2	15.8	25.2	0.42B
A10	60	0.577	17.4	17.0	17.6	22.7	0.20B
B2	30	1.732	29.5	9.6	14.8	29.5	0.40B
B4(1)	60	0.577	14.2	13.9	14.4	18.6	1.72U
(2)	60	0.577	14.8	14.5	15.0	19.3	1.04B
B5(1)	75	0.268	6.9	14.6	14.7	16.5	4.06U
(2)	75	0.268	7.2	15.2	15.3	17.2	0.46B
B6	45	1	24.1	13.6	15.1	24.1	0.94B
B7	15	3.732	28.8	4.4	14.4	28.8	0.38B
B8(1)	45	1	23.7	13.4	14.9	23.7	1.85U
(2)	45	1	24.1	13.6	15.1	24.1	0.86B
B9	45	1	24.0	13.6	15.0	24.0	0.72B
B10	60	0.577	14.8	14.5	15.0	19.5	1.47B
B12	45	1	23.9	13.5	15.0	23.9	0.89B

* B = cracked U = uncracked

Note 1.—The endurances given in Table I represent the total number of reversals sustained by each specimen, to produce the cracks as illustrated in the diagrams and photographs. In many cases they do not represent the endurance at which the crack was first detected.

Note 2.—Position of maximum value of ϕ, σ (and σ_T in case of $\alpha = 45°$), may be obtained by reference to Figs. 2(*a*)-(*c*).

Note 3.—Stresses quoted are average values; at any instant during a test, the stress may have departed by up to about ±10 per cent. from the average value, owing to variations in the frequency of the electricity mains.

†At top or bottom corner of minimum section, *i.e.* at ends of section diagonal perpendicular to neutral plane.

‡At centres of sides of minimum section.

(1), (2)..... indicate 1st, 2nd..... tests on same specimen.

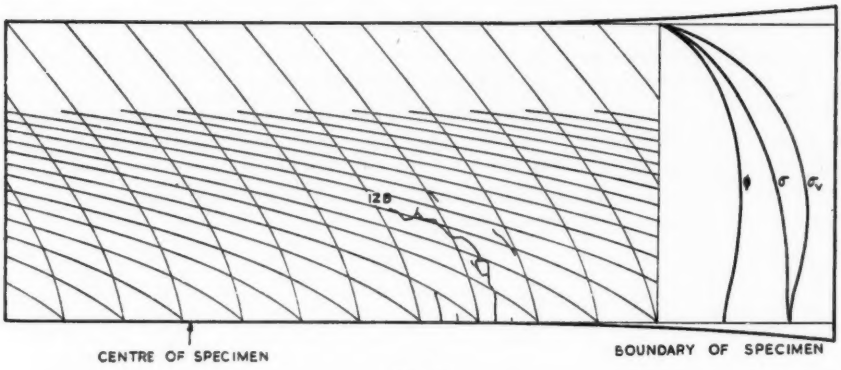

Fig. 6.

Distributions and directions of maximum direct stress and maximum shear stress and courses of fatigue cracks, for $\alpha = 45°$, $f/q = 1$. (Showing cracks on one face of one specimen.)

substantially constant over the central parallel or nearly parallel test section, whereas across the width it falls away from the maximum near the centre of the face to zero at one edge and a low value at the other. One or two of the shear cracks showed a tendency to branch sideways along the transverse plane of maximum shear, but these branches never advanced far.

The regularity of behaviour in this respect of all the pieces cracked in shear gives unexpectedly strong confirmation of the importance of stress gradient, which has been emphasised in other connections[2]. It would appear worth while to explore the point further by extending tests in torsion to wider pieces of rectangular section.

For the value $\alpha = 45°$ fatigue cracks formed sometimes as a result of tension and sometimes as a result of shear. Cracks starting at the edge due to tension are shown in Figs. 5, 6, 7, 16, 17, 18, 22 and 23, and a crack starting in the region of high shear stress and following the shear stress plane is shown in Figs. 8 and 11, and again in Fig. 12. Although two of these cracks are associated with a tensile crack at the edge (Figs. 11 and 12), the crack in Fig. 8 (Specimen A6), which never reached the edge, provides indisputable proof that shear cracking can occur in preference to tensile cracking at this setting. That is, when the ratio of σ_{max} to ϕ_{max} falls to 1.59 failure in shear becomes equally likely with failure in tension.

The vagaries of the testing machine precluded any attempt to observe closely the gradual extension of any crack, but from Figs. 5 to 8 it is clear that cracks starting in tension tend to follow the plane of maximum tension well into regions of high shear and similarly, that cracks starting in shear do not readily deviate into the plane of maximum tension. The tendency to persist on the plane of formation seems to be stronger for a shear crack than a tensile one, tensile cracks persisting down to about $\sigma/\phi = 1.5$ only, whereas shear cracks occasionally persist up to $\sigma/\phi = 1.7$ or

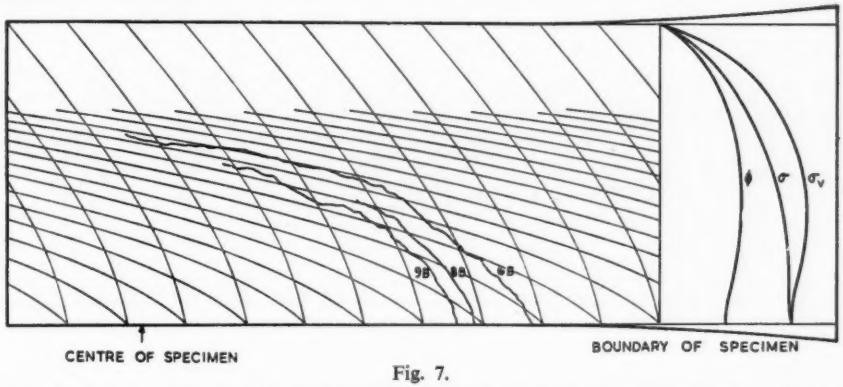

Fig. 7.

Distributions and directions of maximum direct stress and maximum shear stress and courses of fatigue cracks, for $\alpha = 45°$, $f/q = 1$. (Showing crack on one face of each of three specimens.)

even higher (Specimens A6 and A5). However, these observations relate to the stress distribution in the uncracked test piece and the redistribution of stress due to cracking is likely itself to exert considerable influence.

6. General Conclusions

In the present investigation attention has been concentrated on the general directions of propagation of fatigue cracks, and no attempt has been made to examine the detail of this propagation in relation to metallurgical structure, except with a single specimen (B12) (*see* Figs. 22 and 23). The face shown in Fig. 23 was etched and examined in Metallurgy Division, N.P.L. The cracks were found to be

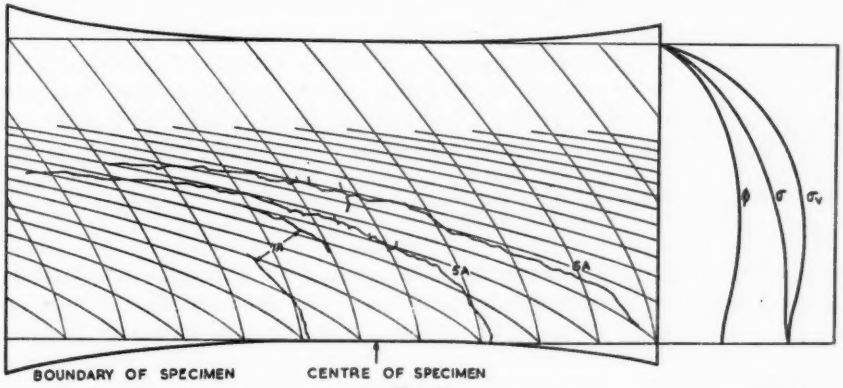

Fig. 8.

Distributions and directions of maximum direct stress and maximum shear stress and courses of fatigue cracks, for $\alpha = 45°$, $f/q = 1$. (Showing cracks on one face of each of three specimens.)

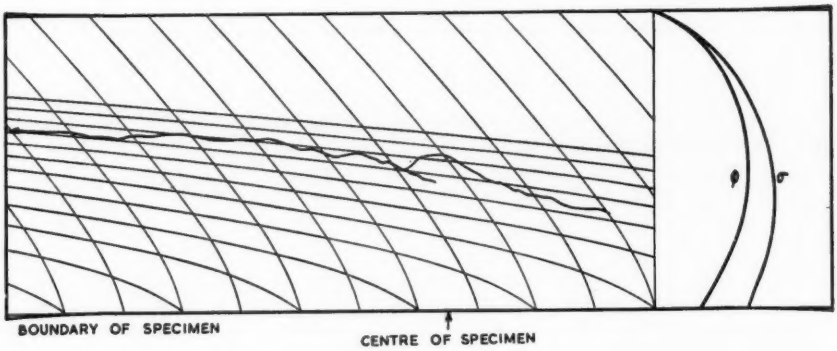

Fig. 9.

Distributions and directions of maximum direct stress and maximum shear stress and course of fatigue crack, for $\alpha=60°$, $f/q=0.577$.

almost entirely trans-crystalline. There was a good deal of branching; at a few points this was quite marked (*see* Figs. 6 and 23), but in these cases the branches reunited after a short distance to continue the general course of the crack, as established fairly definitely for the $\alpha=45°$ case by Figs. 7, 16, 17 and 18.

There was no apparent effect of microstructure which might account for the general trend of the crack, nor for the branches and sudden changes of direction visible under high magnification.

These observations, therefore, although limited, do indicate that the general course of the fatigue crack is influenced only slightly by metallurgical structure. In particular, although the actual initiation of the fatigue crack may be determined by some universal and possibly complex criterion of failure which may be related to the crystal structure, right from the start the crack affords evidence of its origin in direct or shear stresses by propagating at once along the corresponding plane. This conclusion has two separate aspects, both rather surprising; first, that the metallurgical structure, in respect of slip planes, crystal boundaries, and so on, influences the detail of the course of the crack and yet fails to affect its general direction; second, that the initiation of the crack is determined *either* by direct stress *or* by shear and that no single criterion appears at all likely to be able to cover both cases.

If the metallurgical structure had no influence at all, it is conceivable that, by analysis of the stress distribution in the region of the advancing crack, it might be shown that the general direction of propagation along the plane of shear or the plane of direct stress is conditioned by the σ/ϕ ratio in the manner actually observed. But in fact the metallurgical structure does cause the spreading crack to deviate through quite large angles and yet the crack always seems to revert to the "correct" line for some reason as yet unexplained.

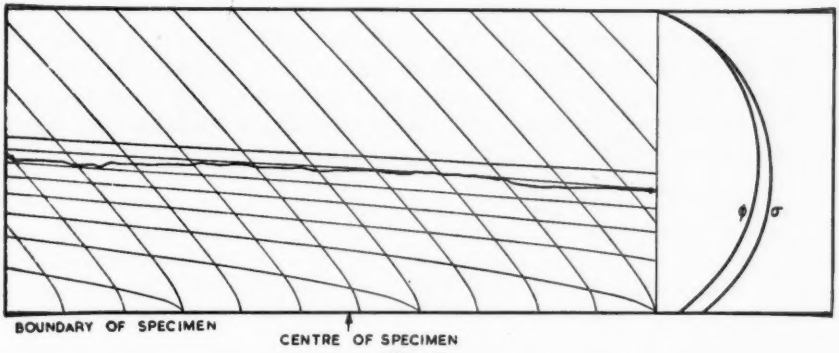

Fig. 10.

Distributions and directions of maximum direct stress and maximum shear stress and course of fatigue crack, for $\alpha=75°$, $f/q=0.268$.

The second point is perhaps yet more startling. Comparison of the course of the crack with the directions of maximum shear and direct stress was undertaken in the first place simply because these directions can be precisely defined, whereas other directions possibly associated with more popular criteria of failure—such as the von Mises-Hencky criterion of shear strain energy—could not be easily represented. From the results, however, it seems that the more elementary comparison is far more appropriate than was originally expected; these two directions predominate and cracking usually occurs *either* where the shear stress is high *or* in the region of high direct stress. For comparison, in Fig. 2*b* the equivalent stress amplitude in accordance with the von Mises-Hencky criterion is plotted on the diagram for $\alpha=45°$; the distribution of this stress amplitude is in fact very little different in form from that of ϕ itself and for other values of α the difference from ϕ is so slight that the extra curve is not worth plotting separately. For $\alpha=45°$ the shear energy criterion gives a single maximum at about $n=0.6$, whereas many of the test pieces cracked at the edge ($n=1$); moreover these edge cracks (*see*, for example, Figs. 5, 6, 7, 16, 17, 18, 22 and 23) definitely followed the plane of maximum direct stress.

In order to account for the positions of the cracks when $\alpha=45°$, any *single* criterion of failure must either afford a flat curve from $n=0.6$ to $n=1$, or preferably, give two equal peaks at these values of n with a dip in between; because in the pieces in general, but particularly in the case of A7, there is strong evidence that the crack starts either at $n=0.6$ or $n=1$ and probably not in between. A double hump curve of this kind could result from a criterion of the form $\phi+\lambda\sigma=$constant, but the value of λ required would appear far too high, and it seems unlikely that such a criterion could be made consistent with the results for higher values of α. In other words $\phi+\lambda\sigma=$constant, with a sufficiently large value of λ, would imply too great a strength in pure torsion in comparison with the strength in bending. It is, in fact, difficult to conceive any single criterion which would account for all the features of the cracks observed in the present series of tests.

Once the conception of a dual alternative is accepted, the dividing line between tensile and shear failure is subject only to the restriction that the critical value of σ/ϕ must lie between 1 and 2. The value indicated by the tests is about 1.6 and at least it lies between 1.3 and 1.7. Further tests are needed to establish the critical condition more certainly.

7. Further Developments

The form of test piece used in the present tests is only one of a number of types which might be used and by appropriate choice of the form, better separation of regions of high shear from those of high direct stress might be achieved. In the limit, development along these lines resolves into independent tests on two or more separate pieces, but before that limit is reached forms of test piece exist in which the whole region between the two regions of maximum shear and maximum direct stress should be available for examination. One suitable form may be a narrow, deep rectangular section subjected to torsion, and to bending in the plane of its depth.

The present exploratory tests on ordinary mild steel drawn from store are to be repeated on materials about which more details of composition and metallurgical structure are available. In particular, tests on aluminium alloy of the type tested in Ref. 3 should be interesting in relation to the previous observations of crack propagation made on this material in the form of thin sheet metal panels.

Acknowledgments

The work described in this paper was done in the Engineering Division of the National Physical Laboratory as part of the research programme of the Mechanical Engineering Research Board. Permission to publish this paper is acknowledged by the authors to the Director of the National Physical Laboratory and the Director of Mechanical Engineering Research of the Department of Scientific and Industrial Research.

References

1. Gough, H. J. (1949). Engineering Steels under Combined Cyclic and Static Stresses. Institution of Mechanical Engineers Presidential Address. 21st Oct. 1949.
2. Phillips, C. E. and Heywood, R. B. (1951). The Size Effect in Fatigue of Plain and Notched Steel Specimens Loaded under Reversed Direct Stress. *Journal of the Institution of Mechanical Engineers.* 23rd Feb. 1951.
3. Phillips, C. E. and Fenner, A. J. (1951). Some Fatigue Tests on Aluminium Alloy and Mild Steel Sheet, with and without Drilled Holes. *Journ. Inst. Mech. Engrs.* Feb. 1951.
4. Gough, H. J., Pollard, H. V. and Clenshaw, W. J. (1951). Some Experiments on the Resistance of Metals to Fatigue under Combined Stresses. R. & M. 2522, 1951.
5. Gough, H. J., Cox, H. L. and Sopwith, D. G. (1934). A Study of the Influence of the Inter-crystalline Boundary on Fatigue Characteristics. *Journal of the Institute of Metals.* Vol. 54, p. 193, 1934.
6. Todhunter, I. (1893). *History of the Theory of Elasticity and Strength of Materials.* Edited by Carl Pearson. Cambridge University Press, 1886-93.
7. Gough, H. J. and Pollard, H. V. (1935). The Strength of Metals under Combined Alternating Stresses. *Proceedings Institution of Mechanical Engineers.* Vol. 131, p. 3, 1935.

The Sensitivity and Range Required in a Toepler Schlieren Apparatus for Photography of High-Speed Air Flow

W. A. MAIR

(Fluid Motion Laboratory, University of Manchester)

Summary: A simple theoretical analysis is given of the quantities affecting the sensitivity and range of a schlieren apparatus. In regions of large density gradient the deflections of the light rays may lie outside the normal working range. When this occurs, dark regions may appear in parts of the photograph that would normally be light. Examples of this effect are given, and it is shown that confusion can be avoided by taking photographs in pairs, with the knife edges or slits arranged to give sensitivities of opposite sign.

The deflections of light rays in various typical disturbances in a supersonic air stream are calculated, using geometrical optics. Prandtl-Meyer expansions, boundary layers, and plane and curved shock waves are considered. From these calculated deflections it is shown that some of the anomalous effects commonly observed in schlieren photographs can be explained.

1. Introduction

The Toepler schlieren apparatus is widely used for investigating high-speed air flow. The apparatus can be used in various forms, as described by Schardin[1], Barnes and Bellinger[2] and other authors.

Holder and North[3] have drawn attention to the importance of the *range* of the system, defined as the range of deflections of a light ray for which the image of the light source or first slit is *partially* cut off by the second knife edge or slit. They showed that if the apparatus is adjusted so that the range is the same for deflections in either direction, the sensitivity is inversely proportional to the range. (The sensitivity is defined as the ratio of the relative change of illumination to the angular deflection of a light ray in a plane normal to the slits.) In using the apparatus to study high-speed air flow it is often necessary to make the sensitivity large enough to show fairly small density gradients. The range is then rather small, and misleading effects may be observed in regions where the density gradients are large.

Paper received March 1951.

[The Aeronautical Quarterly, Volume IV, August 1952]

In this paper the factors affecting the sensitivity and range are considered in some detail. The deflections of rays of light passing through various disturbances in an air stream are calculated, and the effects of exceeding the range of the system are considered.

All the work is based on geometrical optics. It is known that this gives erroneous results, especially when the system is used at high sensitivity[4], but it is believed that a theory of this kind is sufficient to explain qualitatively many of the phenomena that have been observed.

Notation

Suffixes 1 and 2 refer respectively to the points where the light ray enters and leaves the curved shock wave

a	speed of sound
c	radius of curvature of shock wave
D	diameter of field of view
d	width of working section of wind tunnel
e, f, g	*see* Fig. 2
h	height of working section of wind tunnel
I	illumination on screen
I_0	illumination on screen in undisturbed state
$I_{max.}$	maximum illumination on screen
ΔI	change of illumination on screen
K	$=\rho_0 k$ (*see* equation (1))
k	$=(\mu-1)/\rho$
L	length of light path from working section to second slit or knife edge
l	length of light path in shock wave or expansion region
M	Mach number
n	distance in direction of maximum density gradient
R	radius of curvature of light ray
r	distance from centre of Prandtl-Meyer expansion
s	$=\Delta\rho/\rho$, strength of shock wave
t	thickness of shock wave or expansion region
x	in Sections 4 and 5 (*see* Fig. 8)
x, y	rectangular co-ordinates in Section 6 (*see* Fig. 13)
α	inclination of incident light ray to plane shock wave or of tangent at point on curved shock wave to axis of y
$\Delta\alpha$	angle between Mach lines at upstream and downstream boundaries of Prandtl-Meyer expansion
β	mean inclination of light ray to shock wave
γ	ratio of specific heats
ε	angular deflection of light ray
ε'	angular deflection of light ray in plane normal to slits
$\varepsilon_{max.}$	maximum angular deflection of light ray

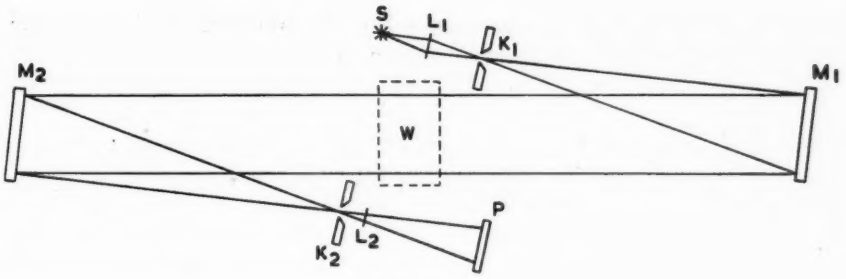

Fig. 1.
Schlieren apparatus.

θ	$=t/c$, ratio of thickness of a shock wave to its radius of curvature
λ	$=\theta/\Delta\mu$, ratio of radius of curvature of light ray to that of shock wave
μ	refractive index
$\Delta\mu$	change of refractive index in shock wave
ν	kinematic viscosity
ρ	density
ρ_0	stagnation density or density of air at normal temperature and pressure
ϕ	angle between slits and plane in which deflection ε is measured

2. Principles

The refractive index μ of a gas is related to the density ρ by the equation

$$\mu-1=k\rho,$$

where k is a constant for a given gas and a given wavelength of light. For air, it is convenient to express this in the form

$$\mu-1=K\frac{\rho}{\rho_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where ρ_0 is the density at normal temperature and pressure. The constant K is dimensionless and varies between 0.000290 and 0.000298 for the visible light spectrum. Equation (1) shows that the gradient of refractive index is directly proportional to the density gradient.

If a parallel beam of light passes through air in which there is a density gradient normal to the direction of the beam, the light travels more slowly where the density is greater and hence the beam is refracted towards the region of greater density.

If the density gradient is uniform it can easily be shown that the light rays are circular arcs whose radius of curvature is given by

$$\frac{1}{R} = \frac{d\mu}{dn} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where $d\mu/dn$ is the gradient of refractive index in a direction normal to the light beam.

In a schlieren apparatus, the angular deflection of each ray of light passing through the air stream is made to produce a corresponding change of illumination on a screen or photographic film. Fig. 1 shows a typical form of the apparatus, using two concave mirrors. (There are many other possible arrangements, but the principles of all of these can be understood by considering the system shown in Fig. 1).

In Fig. 1, light from the source S is focused by a condenser lens L_1 on to a slit K_1. This slit is in the focal plane of the concave mirror M_1, so that a parallel beam of light passes through the wind tunnel W to the concave mirror M_2. A second slit K_2 is placed in the focal plane of the mirror M_2, so that an image of K_1 is formed at K_2. The slits are adjusted so that the image of K_1 is very nearly cut off by one edge of K_2. Then if a density gradient in the wind tunnel causes a refraction of the light beam in any direction that is not parallel to the slits, there will be a change in the amount of light cut off at K_2, and hence a corresponding change in illumination on the screen or film P. The convex lens L_2, in conjunction with the mirror M_2, gives a real image of an object in the wind tunnel on the screen P. Thus the point on the screen P reached by a given ray of light is not displaced by refraction of the ray in the wind tunnel. This is an important feature of the apparatus; it means that all the details of the flow are shown on the screen or photograph in their correct relative positions. (In a simple shadowgraph this is not so).

The apparatus could be used with single knife-edges instead of slits at K_1 and K_2. In many cases, however, it is convenient to use slits, because the sign of the density gradient producing a given change of illumination can then be altered easily by changing the "cut-off" from one edge of the slit to the other. When slits are used, the slit K_1 is usually arranged to restrict the light so that only the part of the source that is of approximately uniform brightness is used. The slit K_2 should be arranged so that only one edge of the slit is illuminated when there are no density gradients causing refraction of the light rays. The other edge of the slit K_2 should be adjusted so that it does not reduce the effective aperture of the object lens L_2 below a reasonable value. It is shown later that the sensitivity and range depend on the dimensions and relative positions of the slits.

Even when single knife-edges are used the system may be nearly equivalent to one using slits. The edge of the object lens L_2 restricts the light passing the second

knife-edge and, if the distance from K_2 to L_2 is small, this gives an effect similar to that produced by a fairly wide slit. In the same way, the finite size of the light source restricts the light passing over the first knife-edge, and gives an effect similar to that produced by a slit.

In the analysis which follows it is assumed that slits are used, as in Fig. 1. It should be emphasised, however, that this analysis applies also to a system using single knife-edges; the widths of the slits considered in the analysis are then the equivalent values corresponding to the size of the light source and the aperture of the object lens L_2.

3. Sensitivity and Range

Figure 2 shows the image of the pair of knife-edges (*PP*) forming the first slit, superposed on the knife-edges (*QQ*) forming the second slit. The dimensions *e*, *f*

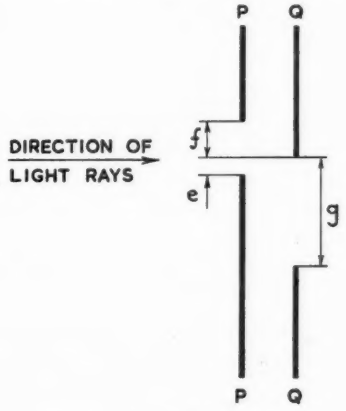

Fig. 2.
Dimensions affecting sensitivity and range.

and *g* refer to the system in its undisturbed state, with no refraction in the wind tunnel. In this state the light is cut off except in a narrow band of width *e*.

Assume that there is a uniform density gradient in the wind tunnel, producing an angular deflection ε in a plane making an angle ϕ with the slits. Then the angular deflection in a plane normal to the slits is

$$\varepsilon' = \varepsilon \sin \phi .$$

If L is the total length of the light path from the working section to the second slit, the corresponding displacement of the image of the first slit (in the plane of the second slit) is $\varepsilon' L$. If the illumination on the screen is I_0 in the undisturbed state,

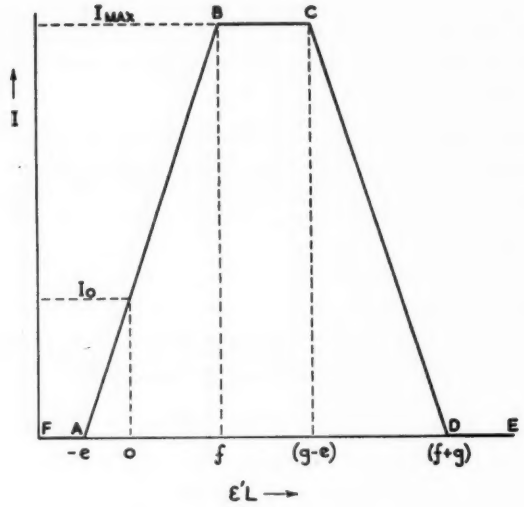

Fig. 3.

Sensitivity and range.

and if this changes by an amount ΔI for a small deflection ε', the *sensitivity* of the system is

$$\frac{\Delta I}{\varepsilon' I_0} = \frac{L}{e} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where the dimension e is as shown in Fig. 2. According to this simple theory the sensitivity can be increased indefinitely by making e sufficiently small. In practice, the maximum sensitivity that can be used is limited by the following considerations.

(*i*) Diffraction effects become increasingly important as e is reduced, so that equation (3) is not correct for very small values of e.

(*ii*) The illumination I_0 decreases with e. However, unless very short exposures are required (one micro-second or less) the light source can be made so bright that this is not a serious limitation.

(*iii*) Optical faults in the apparatus (especially in the glass windows of the wind tunnel) and convection currents in the room cause such disturbances that nothing is gained by increasing the sensitivity beyond a certain amount.

Of these limitations, the last is usually the most important.

Figure 3 shows how the illumination I depends on the displacement of the image of the first slit $\varepsilon' L$. The displacement is considered to be *positive* when its sign is such that the illumination is *increased* (in the normal working range). Then when

$\varepsilon' L = -e$ the light is completely cut off and the illumination is zero (point *A*). When $\varepsilon' L < -e$ the illumination is still zero and the system no longer responds to small changes of ε' (line *AF*). In this state the system is said to be *overloaded*.

Considering now positive values of ε', giving increased illumination in the normal working range, when $\varepsilon' L = f$ all the light from the first slit passes through the second slit, and the illumination has the maximum value $I_{max.}$ (point *B*). (This is only true if $g > (e+f)$, but this condition will usually be satisfied). If $(g-e) > \varepsilon' L > f$ the illumination is the same as for $\varepsilon' L = f$ and again the system is overloaded (line *BC*). Thus the straight line *AB* represents the normal working range of the apparatus, and it is only when $\varepsilon' L$ is within this range $(-e < \varepsilon' L < f)$ that there is a unique value of $\varepsilon' L$ for each value of the illumination *I*.

Holder and North[3] suggested making the range the same for positive values of ε' as for negative values, *i.e.* making $e = f$. If this is done the range is inversely proportional to the sensitivity, for a given value of *L*. Alternatively, it may sometimes be better to choose *e* to give the required sensitivity and then make *f* considerably greater than *e*. In this way an increased range is obtained for positive values of ε', although the range for negative values is still limited by the value of *e* giving the required sensitivity.

Normally, the width of the second slit (*g*) is greater than that of the image of the first slit $(e+f)$, and the range for positive values of ε' is then determined by *f*. If $g < (e+f)$ the maximum illumination is reached for $\varepsilon' L = (g-e)$ and the range for positive values of ε' is then determined by $(g-e)$.

In the normal case where $g > (e+f)$, if $\varepsilon' L > (g-e)$ an increase of $\varepsilon' L$ then gives *less* illumination, as shown by the line *CD* in Fig. 3. If $\varepsilon' L > (f+g)$ there is *zero* illumination (line *DE*).

Holder and North[3] have drawn attention to the importance of the overloaded conditions represented by the lines *AF* and *BC* in Fig. 3. Within the range of density gradients making either $\varepsilon' L < -e$ or $f < \varepsilon' L < (g-e)$ it is impossible to distinguish between different density gradients, so that important details of the flow may sometimes not appear in the photograph. This difficulty can be overcome by increasing *e* and *f* to give increased range, but if *e* is increased too much the sensitivity may become too small for some parts of the field. If a high sensitivity is required in part of the field, two photographs can be taken, one with small *e* to give high sensitivity, the other with large *e* to give large range.

The portions *CD* and *DE* of the curve in Fig. 3 are sometimes very important. They indicate that a large density gradient, of a sign which would give *more* light in the normal working range, may give *less light* or *no light at all* if the system is overloaded. It is important to note, however, that this reversal of effect cannot occur if the density gradient is of a sign giving less light in the normal working range. That is, a region in the photograph that should be dark will still be dark even if the system is overloaded, while a region that should be light may be dark if overloading occurs.

These effects are illustrated in Figs. 4, 5, 6 and 7. Fig. 4 is a straightforward schlieren photograph showing supersonic flow past a two-dimensional wedge at a high incidence. The slits are horizontal and arranged so that increasing air density in a downward direction gives more light. The shock waves and expansions are all shown correctly, because overloading does not occur except in regions which should be dark even in the normal working range. The light line on the upstream side of the expansion region above the wedge represents a weak shock wave caused by the slight bluntness of the leading edge. This has been explained in a paper by Bardsley[5].

Figure 5 is another photograph showing supersonic flow past a wedge. In this case the slits are horizontal and arranged so that increasing air density in an upward direction gives more light. Most of the features of the flow are shown correctly, but there is a dark line in the light fan-shaped region representing the shock wave attached to the lower side of the wedge. (The shock wave appears as a fan-shaped region because the flow is not two-dimensional; the Mach number is reduced in the boundary layers on the glass side-plates). The dark line is caused by overloading of the system. It will be shown later that its position near the downstream edge of the "fan" can be explained.

Figures 6 and 7 are two photographs of the same flow, taken with horizontal slits. In Fig. 6 increasing air density in a downward direction should give more light, while in Fig. 7 the effects should be reversed. (It should be noted that the magnification of the image is much greater in Figs. 6 and 7 than in Figs. 4 and 5; the diameter of the field of view is about 1.5 cm. in the former and about 12 cm. in the latter). In Fig. 6 the shock waves and expansion are shown correctly (as they are in Fig. 4) because overloading does not occur except in regions which should be dark in the normal working range. In Fig. 7, where increasing air density in an upward direction should give more light (in the normal working range), the shock wave below the wedge is shown dark instead of light. Also, there is a dark region near the centre of the Prandtl-Meyer expansion that does not show as a corresponding light region in Fig. 6. Both these effects are caused by overloading of the schlieren system. The weak shock wave on the upstream side of the expansion, as shown in Fig. 4 and discussed by Bardsley[5], also appears in Figs. 6 and 7. The light lines on the left of the black line representing the strong shock wave in Fig. 7 correspond to the curved parts of the shock wave near the glass side-plates. The deflections of the light rays are fairly small in these curved parts of the shock wave, and overloading does not occur.

These examples show that no confusion need arise from the reversal effect if the photographs are always taken in pairs, with the slits set first to give a light region and then a dark region for a given sign of density gradient. Any effect that appears on both photographs cannot be due to overloading and must be a genuine feature of the flow. If a phenomenon appearing in one photograph has no counterpart in the other it should be regarded with suspicion. For example, the dark region near the

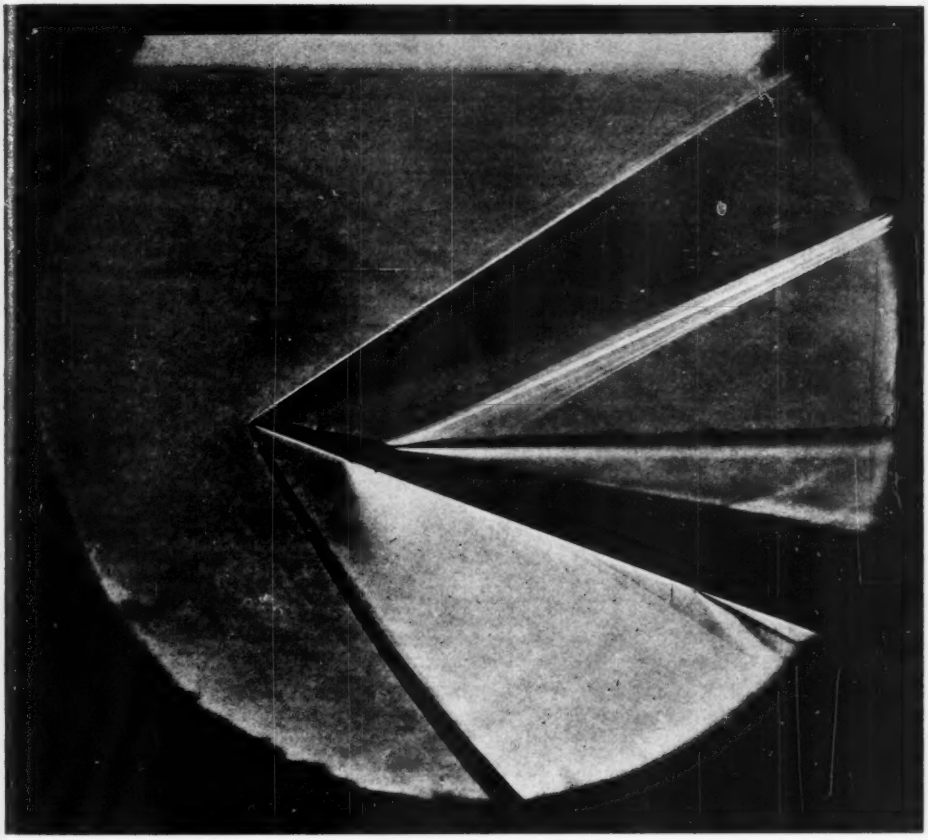

Fig. 4.

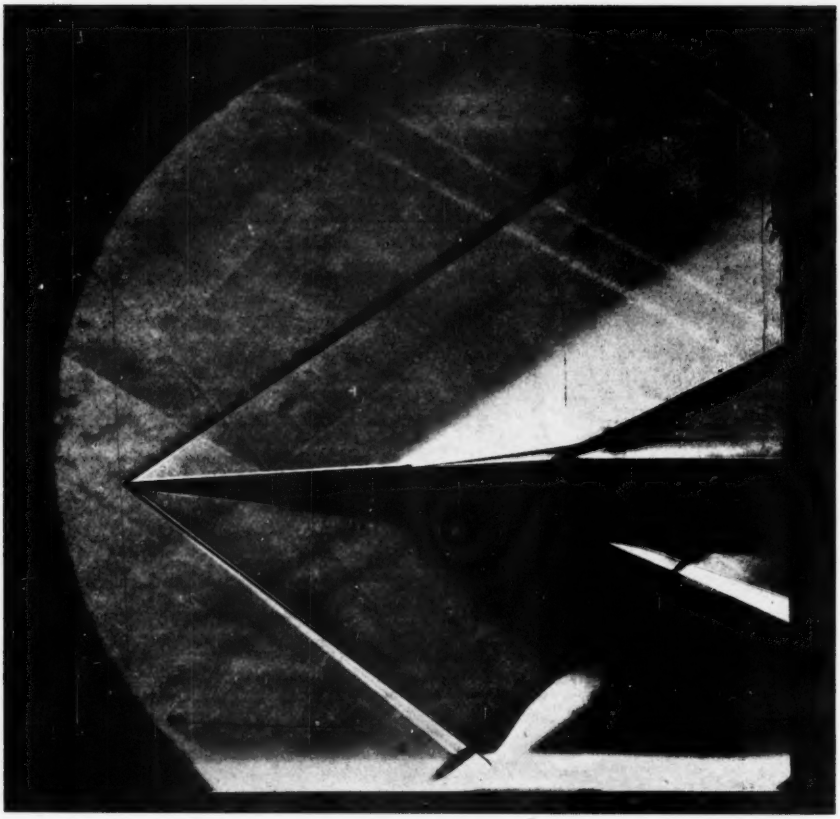

Fig. 5.

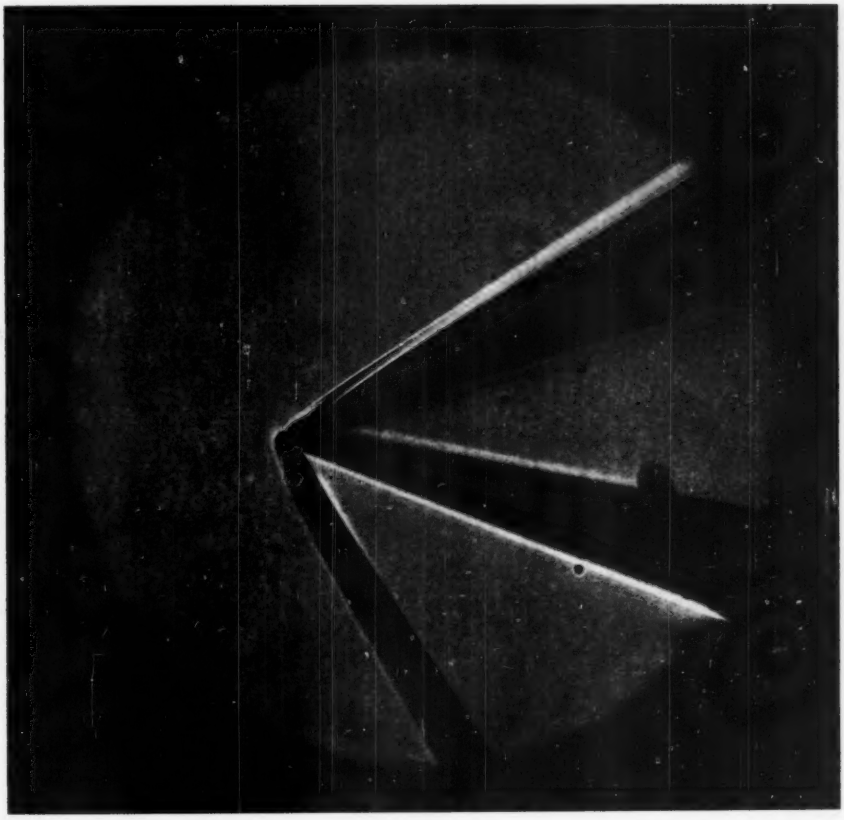

Fig. 6.

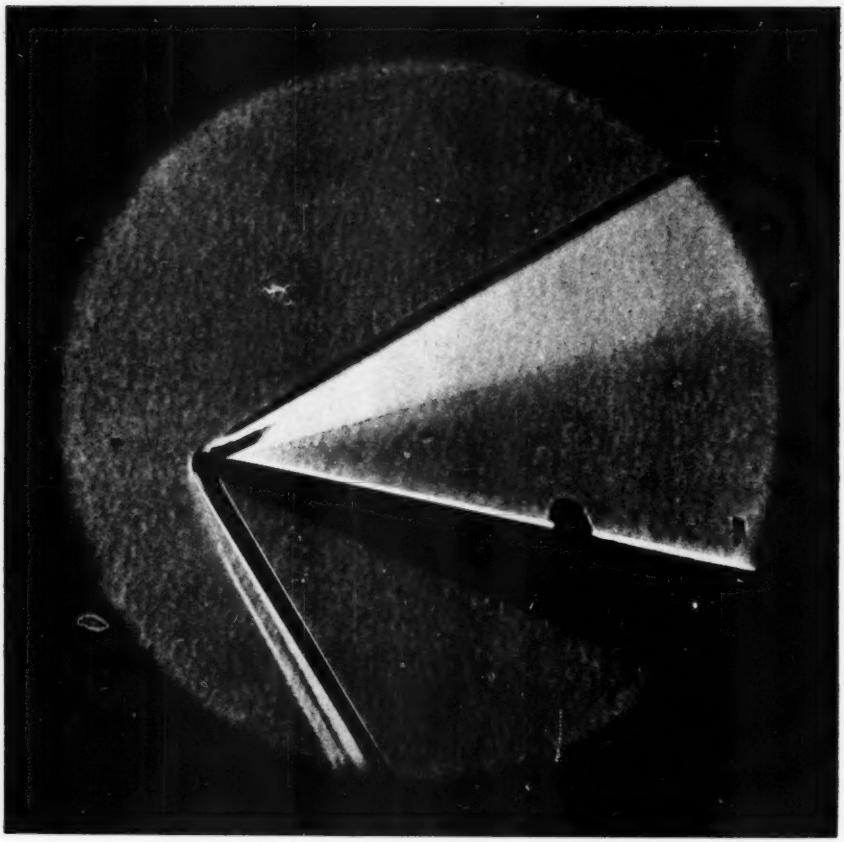

Fig. 7.

centre of the Prandtl-Meyer expansion in Fig. 7 might be interpreted (at first sight) as a shock wave. However, since there is no corresponding light region in Fig. 6 it is certain that there is no shock wave in this position and that the dark region in Fig. 7 is entirely due to overloading of the schlieren system.

It would probably be possible to make the dimension g (Fig. 2) very large compared with e or f and hence to avoid the reversal effect, but this would necessitate the use of a very large aperture lens for the objective L_2, and might introduce troublesome aberrations. Since the reversal effect need not cause any confusion if it is well understood, there is no strong reason for trying to avoid the effect in all cases by making g very large.

Methods of increasing the maximum range that can be obtained with a given light source (*i.e.* increasing e and f) have been discussed by Holder and North[3]. It should be emphasised, however, that an increase of range for deflections giving *less* light necessarily implies a reduction of sensitivity.

In the foregoing discussion it is assumed that the slits are very long in comparison with their width. In practice the slits are often nearly square, so that if there is a strong component of density gradient parallel to the working edge of the slit, reversal effects may occur because light is cut off by one of the "ends" of the slit.

Anomalous effects may also appear near the edge of the field of view, if rays from regions of strong density gradient are refracted so much that they pass outside the second lens or mirror.*

The magnitudes of the deflections ε produced by various disturbances in a supersonic air stream are now considered.

4. Prandtl-Meyer Expansions

A simple example of a weak density gradient of fairly large extent is a Prandtl-Meyer expansion at a large distance from the centre.

If r is distance from the centre of the expansion and n is distance measured normal to the Mach lines, then with $\gamma=1.4$ the density gradient is given by

$$\frac{1}{\rho_0}\frac{d\rho}{dn}=\frac{5}{6r}(M^2-1)^{\frac{1}{2}}\left(1+\frac{M^2}{5}\right)^{-\frac{5}{2}} \quad . \quad . \quad . \quad . \quad (4)$$

where ρ_0 is the stagnation density and M is the local Mach number at the point considered.

Putting $M=2$, with ρ_0 as the density of air at N.T.P., and substituting $K=2.94\times10^{-4}$ in equation (1) gives

$$\frac{d\mu}{dn}=r^{-1}\times10^{-4} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Note added in proof.—An effect of this kind has been observed recently. An oblique shock wave near the edge of the field of view appeared dark instead of light.

If d is the width of the tunnel in the direction of the light beam, equation (2) then gives

$$\varepsilon = \frac{d}{R} = \frac{d}{r} \times 10^{-4} \text{ radians} \quad . \quad . \quad . \quad . \quad (6)$$

In a tunnel of approximately square section, it will often be necessary to detect expansions of this kind for $r \simeq d$. (In Fig. 4, $r \simeq 0.9\,d$ where the expansion meets the edge of the field of view). The condition $r = d$ gives

$$\varepsilon = 10^{-4} \text{ radian}$$

for this example. This gives an indication of the order of magnitude of the smallest values of ε that it may be required to detect. Fig. 4, taken in conditions corresponding roughly to the values assumed in this example, shows that a deflection of this

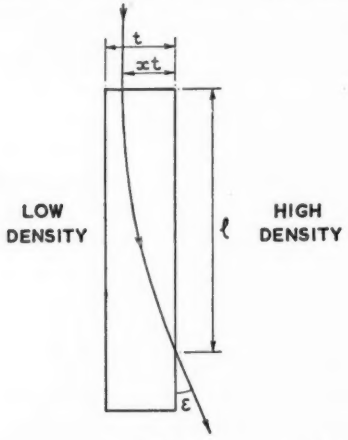

Fig. 8.

Refraction in a Prandtl-Meyer expansion or plane shock wave.

amount can, in fact, be detected quite easily. (The sensitivity used for Fig. 4 was considerably less than the maximum possible).

In the foregoing analysis it is assumed that the value of r is fairly large, so that most of the light rays entering the expansion remain inside it until they reach the other side of the tunnel. If r is small this may not be so, and the effect of the finite extent of the expansion region must be considered.

Let $\Delta\alpha$ be the angle between the Mach lines forming the boundaries of the expansion "fan." Then the thickness of the expansion region in a direction normal to the Mach lines and to the light beam is $t = r\Delta\alpha$. Let a light ray enter the tunnel at a distance xt from the high-density side of the expansion, and emerge from the

expansion after travelling a distance l (Fig. 8). (It is assumed here that the width of the tunnel d is greater than l; if this is not so the deflection ε is given by equations (1), (2) and (4) as described earlier.

For $d > l$,

$$l = R\varepsilon = \frac{2xt}{\varepsilon}.$$

Therefore
$$\varepsilon^2 = \frac{2xt}{R} = \frac{2xr\Delta\alpha}{R} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Since $R \propto r$, for given values of ρ_0 and M, this leads to

$$\varepsilon^2 \propto x\Delta\alpha \text{ (independent of } r).$$

For the conditions assumed in the numerical example,

$$\frac{R}{r} = 10^4,$$

and therefore
$$\varepsilon = (2x\Delta\alpha)^{\frac{1}{2}} \times 10^{-2} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

and the maximum deflection (for $x = 1$) is

$$\varepsilon_{max.} = (2\Delta\alpha)^{\frac{1}{2}} \times 10^{-2} \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Thus for this example the maximum deflection is given by equation (6) or (9), whichever gives the *smaller* value of $\varepsilon_{max.}$. For large values of r the length of the light path is limited by the width of the tunnel and $\varepsilon \propto r^{-1}$ (equation (6)). For small values of r the length of the light path is limited by the distance $r\Delta\alpha$, and ε is independent of r (equation (9)). The critical value of r, separating the two regimes, occurs when

$$d = l = \frac{2xr\Delta\alpha}{\varepsilon} \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

For this example, with $x = 1$, this gives

$$\frac{d}{r} = 100\,(2\Delta\alpha)^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad . \quad (11)$$

In Fig. 7 the conditions are nearly the same as those assumed in the example. $\Delta\alpha$ is about 16° and the value of r at the outer edge of the dark region near the centre of the expansion is about 1½ mm. Equation (9) shows that the maximum deflection should be 0.75×10^{-2} radians and, since $d = 10$ cm., this should occur for all values of r smaller than 1.3 mm. (equation (11)). This deflection is about 100 times the

smallest deflection that can be detected, and is quite large enough to make $\varepsilon' L > (f+g)$ (Fig. 3), thus explaining the dark region near the centre of the expansion in Fig. 7. As already mentioned, no such effect is shown in Fig. 6, where the response to density gradients in the normal working range is of opposite sign.

5. Plane Shock Waves

In calculating the deflection of a light ray passing through a shock wave, the problem will be simplified by assuming that the change of density takes place linearly in a thickness t. It will be shown later that for strong shock waves the deflection for a given strength is independent of t unless the shock wave is highly curved. However, for weak shock waves the deflection may depend on t, and this Section therefore starts with a discussion of the thickness of a weak shock wave.

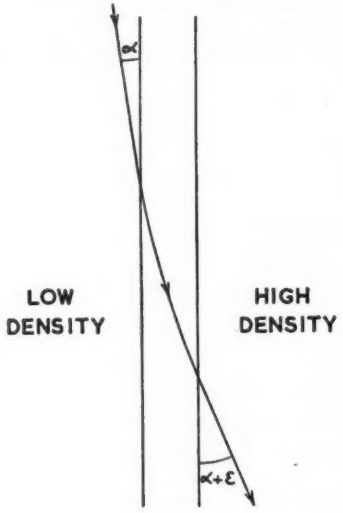

Fig. 9.

Refraction in a plane shock wave with light entering on low-density side.

Taylor's approximate theory[6] shows that for a weak shock wave in air ($\gamma = 1.4$, Prandtl number $= 0.75$) the thickness in which 90 per cent. of the total change of density occurs is

$$t = \frac{9.3\nu}{sa} \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

where $s = \Delta\rho/\rho$ and a is the speed of sound at the shock wave. Various other definitions based on the same theory have been considered. For example, with a linear change of density, the thickness for which the assumed density gradient would be equal to the true maximum gradient is 67 per cent. of the thickness given by

equation (12). The thickness in which 98 per cent. of the total change occurs is 50 per cent. greater than the value given by equation (12). For the present purpose the exact choice of definition of t is not very important and the value given by equation (12) will be used in all the calculations.

For air flowing in a wind tunnel at a Mach number of 2, with atmospheric stagnation conditions, equation (12) gives

$$t = \frac{1.5 \times 10^{-4}}{s} \text{cm}. \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

This should be approximately correct for values of s up to about one, but it is known that Taylor's approximate theory underestimates the thickness of a strong shock wave.

Now consider the refraction of a light ray in a plane shock wave, with a linear change of density (and refractive index) in a thickness t. Assume at first that the shock wave is of large extent, so that a ray of light entering the wave at any point emerges again before reaching a boundary of the flow. Let the difference of refractive index across the shock wave be $\Delta\mu$ and let the incident light ray make an angle α with the plane of the wave. Then if α is small the radius of curvature of the light ray inside the wave is

$$R = \frac{t}{\Delta\mu} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

If the light ray enters on the low-density side of the shock wave, the refraction is as shown in Fig. 9. The length of the light path inside the shock wave is

$$l = \frac{t}{\alpha + \frac{\varepsilon}{2}} = R\varepsilon = \frac{\varepsilon t}{\Delta\mu}$$

(assuming α and ε are small), and therefore

$$\varepsilon\left(\alpha + \frac{\varepsilon}{2}\right) = \Delta\mu \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

and the positive root is

$$\varepsilon = (\alpha^2 + 2\Delta\mu)^{\frac{1}{2}} - \alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

When $\alpha = 0$, $\varepsilon = (2\Delta\mu)^{\frac{1}{2}}$

and when $\alpha >> \varepsilon$, $\varepsilon = \frac{\Delta\mu}{\alpha}$

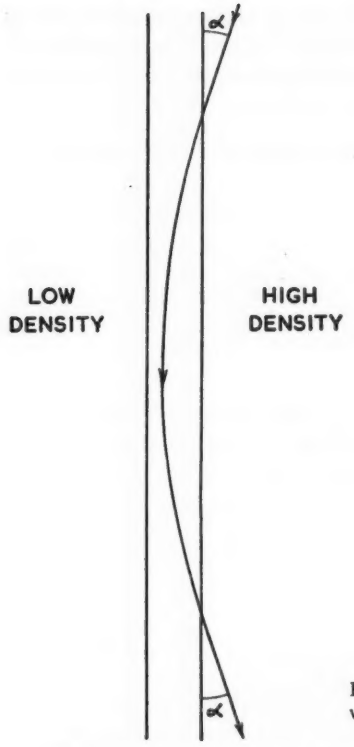

Fig. 10 (*left*).
Refraction in a plane shock wave with light entering on high-density side ($\alpha^2<2\Delta\mu$).

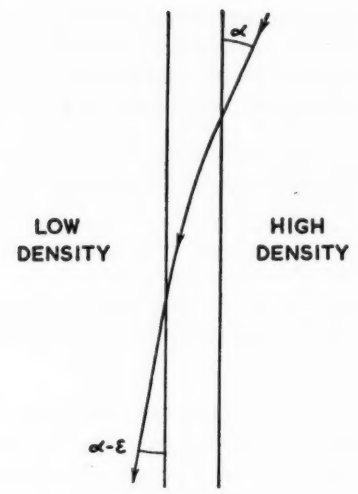

Fig. 11 (*above*).
Refraction in a plane shock wave with light entering on high-density side ($\alpha^2>2\Delta\mu$).

If the ray enters the shock wave on the high-density side at a small value of α, the refraction is as shown in Fig. 10. This type of refraction occurs if the ray does not emerge from the shock wave on the low-density side. The condition for this is

$$R(1-\cos\alpha)<t \text{ where } R=\frac{t}{\Delta\mu},$$

i.e.

$$\alpha^2<2\Delta\mu \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

The shock wave then acts as a *plane mirror*, and the angular deflection of the ray is

$$\varepsilon=2\alpha. \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

If $\alpha^2>2\Delta\mu$, with the ray entering the shock wave on the high-density side, the refraction is as shown in Fig. 11. Then the length of the light path inside the shock wave is

$$l=\frac{t}{\alpha-\frac{\varepsilon}{2}}=R\varepsilon=\frac{\varepsilon t}{\Delta\mu},$$

and therefore $$\varepsilon\left(\alpha-\frac{\varepsilon}{2}\right)=\Delta\mu. \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

Since ε cannot be greater than α, the appropriate root of the equation is

$$\varepsilon=\alpha-(\alpha^2-2\Delta\mu)^{\frac{1}{2}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

and when $\alpha>>\varepsilon$ the solution is

$$\varepsilon=\frac{\Delta\mu}{\alpha}.$$

For the particular case when α^2 is *just greater* than $2\Delta\mu$, we have $\varepsilon=\alpha$ and $d\varepsilon/d\alpha=-\infty$. When α^2 is *just less* than $2\Delta\mu$ the refraction is as shown in Fig. 10 and $\varepsilon=2\alpha$. Thus ε decreases suddenly from $2(2\Delta\mu)^{\frac{1}{2}}$ to $(2\Delta\mu)^{\frac{1}{2}}$ as α increases through the critical value $(2\Delta\mu)^{\frac{1}{2}}$.

To illustrate the significance of these results some calculations have been made for various shock-wave strengths and angles of incidence of the light beam. It has been assumed that the shock waves are formed in a wind tunnel at a Mach number of 2, with normal atmospheric stagnation conditions. For these conditions,

$$\Delta\mu=6.7\,s\times10^{-5},$$

where $s=\Delta\rho/\rho$, the strength of the shock wave.

The full lines in Fig. 12 show the results of these calculations, for values of s between 10^{-3} and one. Unless the wind tunnel is large, the deflections calculated for the weaker shock waves require modification to allow for the thickness of the shock wave and the finite width of the tunnel. This is now considered.

If a ray of light, entering the shock wave near one side of the wind tunnel, emerges from the wave before reaching the other side of the tunnel, the deflection is given correctly by equation (16), (18) or (20). If the width of the tunnel (d) is so small that the ray remains inside the wave for the whole width of the tunnel, the deflection is simply

$$\varepsilon=\frac{d}{R}=\frac{\Delta\mu}{t}d \quad . \quad . \quad . \quad . \quad . \quad . \quad (21)$$

In general, the deflection will be given either by this equation or by the appropriate one of equations (16), (18) or (20), whichever gives the *smaller* value of ε.

Introducing the thickness t from equation (12) and putting

$$\Delta\mu=K\left(\frac{\rho}{\rho_0}\right)s \qquad \text{(from equation (1))}$$

gives $$\varepsilon=\frac{Ka\left(\frac{\rho}{\rho_0}\right)s^2d}{9.3\,\nu} \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

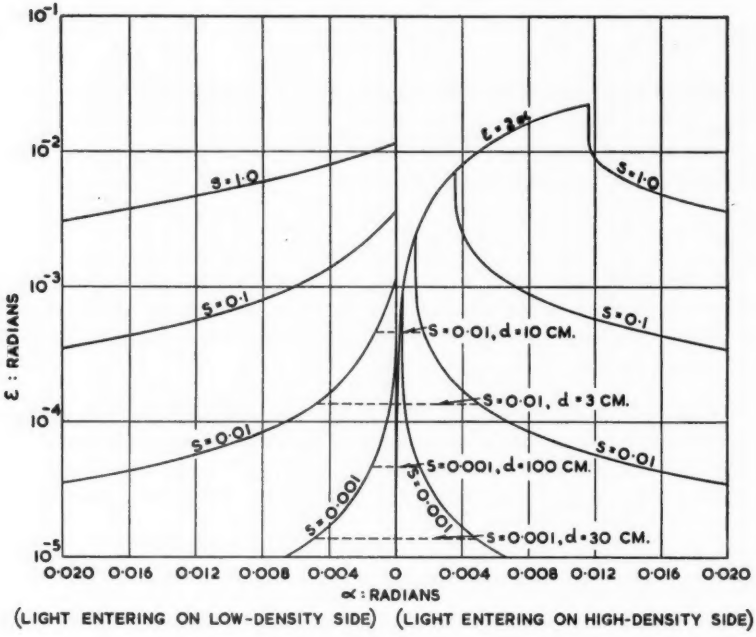

Fig. 12.

Deflections of light rays in plane shock waves of various strengths.

For the conditions assumed in Fig. 12 this becomes

$$\varepsilon = 0.46\, s^2 d \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

It is interesting to note that equation (22) gives a deflection proportional to s^2, whereas the deflection given by equation (16) or (20) is roughly proportional to s, for values of α that are not too small.

The horizontal dotted lines in Fig. 12 show the deflections given by equation (23), where this is applicable, for shock-wave strengths of 0.01 and 0.001 and various values of the tunnel width d. If the tunnel width is such that equation (23) is applicable, the full-line curve above the appropriate dotted line should be ignored. It can be seen that for shock-wave strengths as small as 0.001, the deflections for small values of α are considerably affected by the width of the tunnel, unless this is very large. (It may be noted that for the conditions assumed in Fig. 12, equation (13) shows that $t = 1.5$ mm. for $s = 0.001$).

A diagram such as Fig. 12 can be used to determine the order of magnitude of the smallest shock-wave strength that can be detected. For example, if observations

of Prandtl-Meyer expansions show that a deflection of 10^{-4} radian can be detected, the corresponding shock-wave strength is about 0.005, for $d=10$ cm. and the conditions assumed in Fig. 12. An estimate of this kind may sometimes be unreliable, however, because the optical system may not show up a very *thin* line, representing a plane shock wave, unless the change of illumination is much larger than the minimum that could be detected in a more extended region.

Figure 12 also shows that in a strong shock wave ($s \approx 1$) the deflection may be as much as 0.01 radian. With a length L of several metres this gives a displacement εL of several centimetres, and explains the serious overloading that often occurs in strong shock waves (*e.g.* Fig. 7).

In the special case when $\alpha=0$ the refraction is similar to that discussed earlier for a Prandtl-Meyer expansion (Fig. 8). If the light ray enters the shock wave at a distance xt from the high-density side, and emerges from the wave after travelling a distance l, then

$$l=\frac{2xt}{\varepsilon}=R\varepsilon=\frac{\varepsilon t}{\Delta\mu},$$

and therefore

$$\varepsilon=(2x\Delta\mu)^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

This explains the discontinuity found before at $\alpha=0$ (Fig. 12). As x varies from 0 to 1, ε varies from 0 to $(2\Delta\mu)^{\frac{1}{2}}$, these being the values given by equations (18) and (16) for $\alpha=0$.

Equation (24) is valid if the ray emerges from the shock wave before reaching the other side of the tunnel. If the width of the tunnel is too small to allow this, the deflection is given by equation (22). In general, the deflection is given by equation (22) or (24), whichever gives the *smaller* value of ε.

The special case of a plane shock wave with $\alpha=0$ is not of much importance in practice, because for a strong shock wave the curvature of the wave is usually important, and for a weak wave the deflection is given by equation (22) and is therefore constant for an appreciable range of α, unless the wind tunnel is very large (Fig. 12).

The beam of light entering the wind tunnel can never be composed of rays that are all exactly parallel. This is partly because of the imperfections of the mirror or lens used to produce this beam, and partly because a light source of finite size must be used. The height of the source is usually about 1/5,000 to 1/1,000 of the focal length of the lens or mirror, depending on the sensitivity and range of the system. Taking the greater of these two values, it is found that different rays of light in the approximately parallel beam may be inclined to one another at angles up to about 0.001 radian. Fig. 12 shows that variations in α of this amount will not have any important effect, except for the case of a very weak shock wave in a large wind tunnel, with a very small mean value of α.

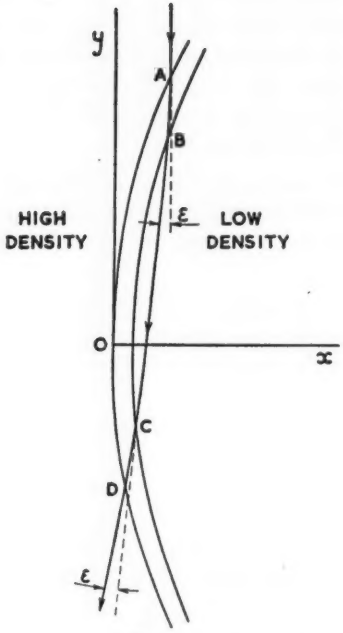

Fig. 13.
Refraction in a curved shock wave
(convex on high-density side)
$\alpha_1^2 > \frac{2\theta}{\lambda}(1+\lambda)$

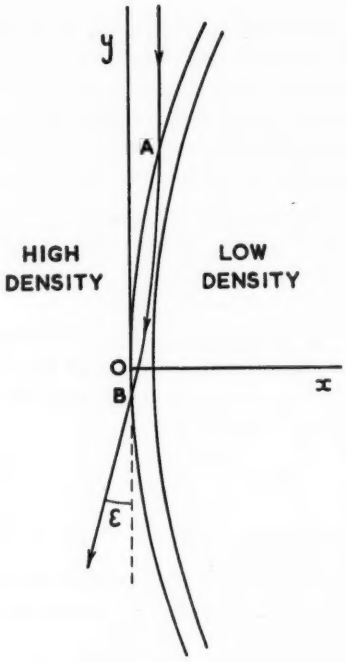

Fig. 14.
Refraction in a curved shock wave
(convex on high-density side)
$\alpha_1^2 < \frac{2\theta}{\lambda}(1+\lambda)$

6. Curved Shock Waves

Schlieren systems are normally used with supersonic wind tunnels only for studying flow that is either axially symmetrical or approximately two-dimensional. Even in two-dimensional flow the shock waves are usually slightly curved, because the Mach number is not exactly constant across the whole width of the working section. In particular, there is a reduction of Mach number, and hence usually a change of shock-wave angle, in the boundary layers on the glass side-plates.

The exact form of a curved shock wave in an approximately two-dimensional flow depends on the shock-wave strength and on the distribution of Mach number across the working section. In most cases the Mach number is higher at the centre of the tunnel than near the edges, so that the shock wave is convex on the high-density side. In order to investigate the refraction of a light ray in such a shock wave, consider a shock wave whose section on the convex side, in a plane normal to the wave and parallel to the light beam, is the parabola $y^2=2cx$ (Fig. 13). Since

only a small part of this parabola is being considered, the radius of curvature may be assumed to be equal to c at all points.

Consider a ray of light entering the shock wave at A in a direction parallel to the y-axis. If A is not too near the origin O the ray will pass twice through the shock wave, emerging finally at D. The system is symmetrical about the perpendicular bisector of BC; hence the deflection ε at AB is equal to that at CD and the total deflection is 2ε.

Let the co-ordinates of A and D be (x_1, y_1) and (x_2, y_2) respectively. Also let

$$\alpha = \frac{y}{c} = \frac{dx}{dy}$$

$$\theta = \frac{t}{c}$$

$$\lambda = \frac{t}{c\Delta\mu} = \frac{\theta}{\Delta\mu}$$

$$\left(\text{then } \Delta\mu = \frac{\theta}{\lambda}\right).$$

It may be noted that λ is the ratio of the radius of curvature of the light ray to that of the shock wave.

Let β be the mean inclination of the light ray to the shock wave at AB or CD. Equating the length of the light path in the shock wave to t/β and to $t\varepsilon/\Delta\mu$, and allowing for the change of slope of the shock wave between A and B, gives

$$\varepsilon^2(1+\lambda) - 2\alpha_1\varepsilon + 2\frac{\theta}{\lambda} = 0 \qquad . \quad . \quad . \quad . \quad (25)$$

where $$\alpha_1 = \frac{y_1}{c} .$$

The appropriate root of this equation is

$$\varepsilon = \frac{\alpha_1 - \left[\alpha_1^2 - \frac{2\theta(1+\lambda)}{\lambda}\right]^{\frac{1}{2}}}{1+\lambda} \qquad . \quad . \quad . \quad . \quad (26)$$

and the total deflection is 2ε.

Because the system is symmetrical about the perpendicular bisector of BC, we also have

$$\alpha_1 + \alpha_2 = 2\varepsilon \qquad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

where $$\alpha_2 = \frac{y_2}{c} .$$

If $\alpha_1^2 < 2\theta(1+\lambda)/\lambda$, equation (25) has no real roots. In this case the ray of light never emerges on the low-density side of the shock wave, and the refraction is

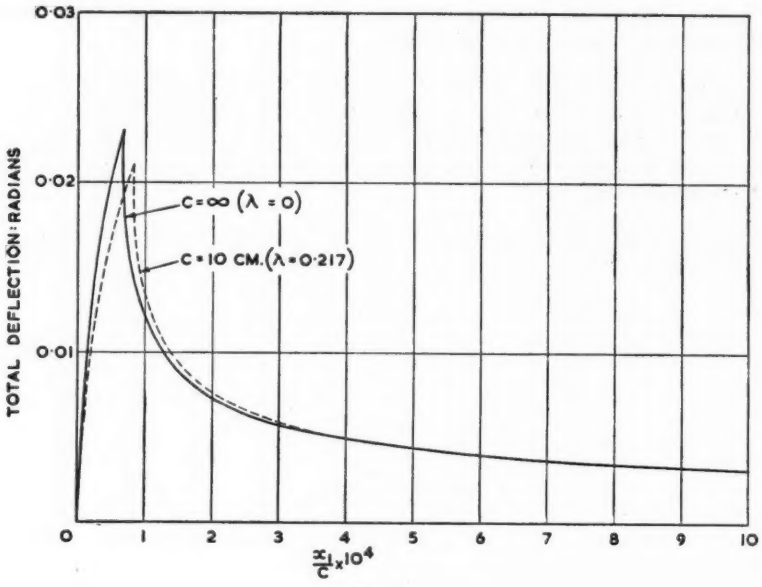

Fig. 15.

Refraction in a curved shock wave (convex on high-density side), $s=\Delta\rho/\rho=1$.

as shown in Fig. 14. Since the light ray and the shock wave are both approximately circular arcs, the inclination of the ray to the shock wave is the same at A as at B, and

$$\varepsilon=\alpha_1+\alpha_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)$$

where suffixes 1 and 2 refer to the points A and B. Also, the length of the light path in the shock wave is

$$y_1-y_2=\frac{t\varepsilon}{\Delta\mu}.$$

Therefore

$$\varepsilon=\frac{2\alpha_1}{1+\lambda} \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

and

$$\frac{y_2}{y_1}=\frac{\alpha_2}{\alpha_1}=\frac{1-\lambda}{1+\lambda} \quad . \quad . \quad . \quad . \quad . \quad (30)$$

When $\lambda=1$, the radius of curvature of the light ray is equal to that of the shock wave, and the point B is at the origin O. When $\lambda>1$, y_2 is negative and when $\lambda<1$, y_2 is positive.

It can easily be verified that when $\alpha_1^2 = 2\theta(1+\lambda)/\lambda$ the light ray just touches the low-density side of the shock wave. The deflection is then

$$\frac{2\alpha_1}{1+\lambda} = 2\left(\frac{2\Delta\mu}{1+\lambda}\right)^{\frac{1}{2}}$$

as found from either of the equations (26) or (29), remembering that with equation (26) the total deflection is 2ε.

If the shock wave has a radius of curvature c that is very large compared with its thickness t, $\theta = 0$, but $\theta/\lambda = \Delta\mu$, and hence $\lambda = 0$ unless $\Delta\mu = 0$. Then equations (25), (26) and (29) reduce to the equations (19), (20) and (18) that were derived earlier for a plane shock wave.

To indicate the appearance of a curved shock wave of this kind in a schlieren photograph, the total deflection is plotted against x_1/c in Figs. 15 and 16, for two strengths of shock wave. All the calculations were for shock waves in a wind tunnel at a Mach number of 2, with atmospheric stagnation conditions. For the stronger shock wave ($s = 1$, Fig. 15), λ is fairly small for all values of c greater than about 10 cm., so that for these values of c the deflection is nearly the same as for $c = \infty$ ($\lambda = 0$). For the weaker shock wave ($s = 0.1$, Fig. 16) the radii of curvature of the light rays are greater, and λ is large enough to have an important effect for values of c up to several metres.

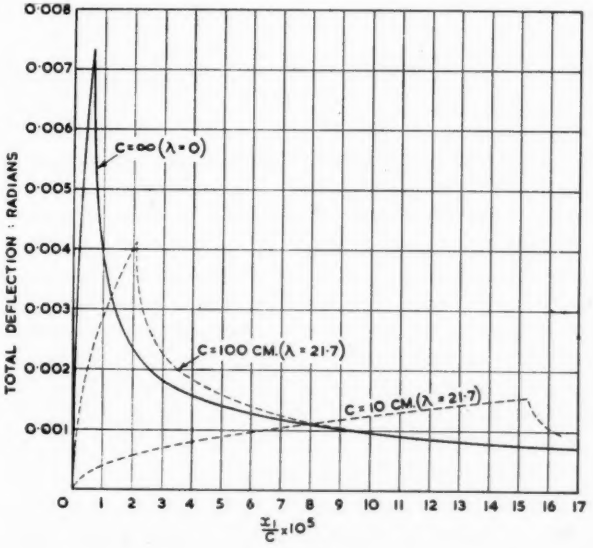

Fig. 16.

Refraction in a curved shock wave (convex on high-density side), $s = \Delta\rho/\rho = 0.1$.

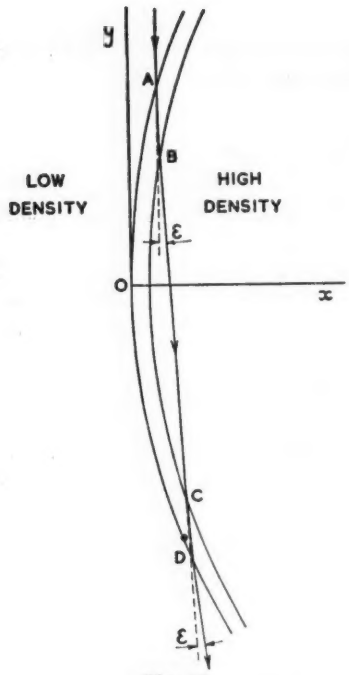

Fig. 17.
Refraction in a curved shock wave
(convex on low-density side)
$x_1^2 > \frac{2\theta}{\lambda}(\lambda - 1)$

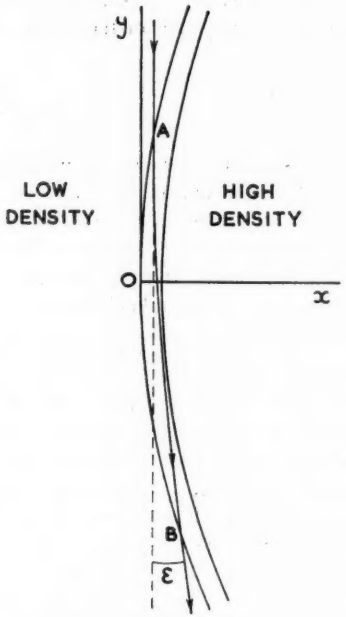

Fig. 18.
Refraction in a curved shock wave
(convex on low-density side)
$x_1^2 < \frac{2\theta}{\lambda}(\lambda - 1)$

A nearly two-dimensional shock wave of the kind considered in Figs. 13 to 16 usually appears in a schlieren photograph as a fan-shaped region (*e.g.* Figs 4 and 5). If the shock wave is convex on the high-density side, the downstream edge of this "fan" represents the central part of the shock wave, *i.e.* the part near the origin in Figs. 13 and 14. Figs. 15 and 16 show that the maximum deflection of a light ray occurs for a small positive value of x_1, corresponding to a position on the photograph near the downstream edge of the "fan." Hence the greatest overloading of the schlieren system occurs at this position. This explains the dark line in Fig. 5, near the downstream edge of the light fan-shaped region representing the shock wave attached to the lower side of the wedge. Figs. 15 and 16 show that for the conditions assumed the maximum deflection is of the order of 10^{-2} radian, a value quite large enough to cause serious overloading of the schlieren system.

In axially symmetrical flow the shock waves are convex on the *low-density* side. This case can be investigated by methods similar to those already used. Again, the convex side of the shock wave can be represented by the equation $y^2 = 2cx$, since only a small part of a shock wave of circular section is being considered.

If the ray of light emerges on the high-density side of the wave the refraction is as shown in Fig. 17. With the same notation and method as were used in discussing Fig. 13,

$$\varepsilon^2(1-\lambda)+2\alpha_1\varepsilon-\frac{2\theta}{\lambda}=0 \quad . \quad . \quad . \quad . \quad (31)$$

and the appropriate root is

$$\varepsilon=\frac{\left[\alpha_1^2+2\theta\,\frac{1-\lambda}{\lambda}\right]^{\frac{1}{2}}-\alpha_1}{1-\lambda} \quad . \quad . \quad . \quad . \quad (32)$$

As before, the system is symmetrical about the perpendicular bisector of *BC*, and the total deflection is 2ε, where ε is given by equation (32). Also

$$\alpha_1+\alpha_2=-2\varepsilon \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (33)$$

For the special case when $\lambda=1$, equation (31) gives

$$\varepsilon=\frac{\theta}{\alpha_1}=\frac{\Delta\mu}{\alpha_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (34)$$

For this case ($\lambda=1$) the radius of curvature of the light ray is equal to that of the shock wave. For $\alpha_1=0$, $\varepsilon=\infty$ and the light ray remains inside the shock wave indefinitely (with the assumptions made in this simplified theory).

Equation (32) gives the deflection for $\lambda<1$, or for $\lambda>1$ and $\alpha_1^2>2\theta(\lambda-1)/\lambda$.

When $\lambda>1$ and $\alpha_1^2<2\theta\dfrac{\lambda-1}{\lambda}$

equation (31) has no real roots. In this case the ray never emerges on the high-density side of the shock wave and the refraction is as shown in Fig. 18.

Using a method similar to that used in discussing Fig. 14,

$$\varepsilon=-(\alpha_1+\alpha_2)=\frac{2\alpha_1}{\lambda-1} \quad . \quad . \quad . \quad . \quad . \quad (35)$$

and

$$\frac{y_2}{y_1}=\frac{\alpha_2}{\alpha_1}=-\frac{\lambda+1}{\lambda-1} \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)$$

Since for this case $\lambda>1$, y_2 is always negative (for y_1 positive).

When $\alpha_1^2=2\theta(\lambda-1)/\lambda$, the light ray just touches the high-density side of the shock wave. Equation (32) or (35) then shows that the deflection is

$$\frac{2\alpha_1}{\lambda-1}=2\left[\frac{2\Delta\mu}{\lambda-1}\right]^{\frac{1}{2}}$$

(When equation (32) is applicable the total deflection is 2ε).

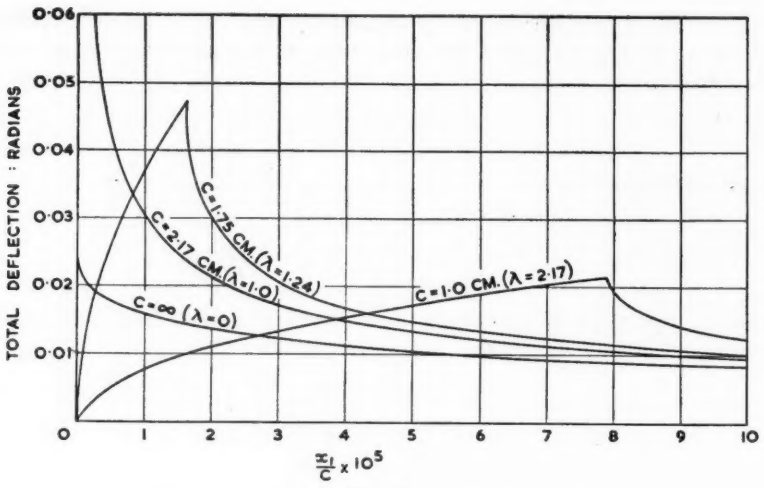

Fig. 19.

Refraction in a curved shock wave (convex on low-density side), $s = \Delta\rho/\rho = 1$.

When c is very large compared with t, $\theta = \lambda = 0$ (unless $\Delta\mu = 0$), and equations (31) and (32) reduce to the equations (15) and (16) derived earlier for a plane shock wave.

Figures 19 and 20 show the results of some numerical calculations on shock waves that are convex on the low-density side. The results are plotted in the same way as in Figs. 15 and 16 and, as before, the calculations are made for shock waves in a wind tunnel at a Mach number of 2, with atmospheric stagnation conditions. For weak shock waves (*e.g.* $s = 0.1$, Fig. 20) λ is greater than 1 for all values of c likely to occur in a wind tunnel with axially symmetrical flow. The shape of the deflection curve (Fig. 20) is then generally similar to the shape found for waves that are convex on the high-density side (Figs. 15 and 16). For strong shock waves (*e.g.* $s = 1$, Fig. 19) the critical case ($\lambda = 1$) occurs for a fairly small value of c. For greater values of c ($\lambda < 1$) the maximum deflection occurs for $x_1 = 0$, in contrast to the curves for $\lambda > 1$.

Figures 19 and 20 show that deflections of the light beam large enough to cause overloading of the schlieren system would be expected for moderately strong shock waves in axially symmetrical flow, but the photographs that have been published show little evidence of this. The probable explanation is that the radius of curvature c is fairly small in most cases, so that the range of values of x_1 giving a large deflection is very small indeed. For example, if $c = 1$ cm., Fig. 19 shows that for $s = 1$ the deflection exceeds 0.01 radian only for values of x_1 between 1.7×10^{-5} and 14×10^{-5} cm. Thus this large deflection occurs only for a band of width 12×10^{-5}

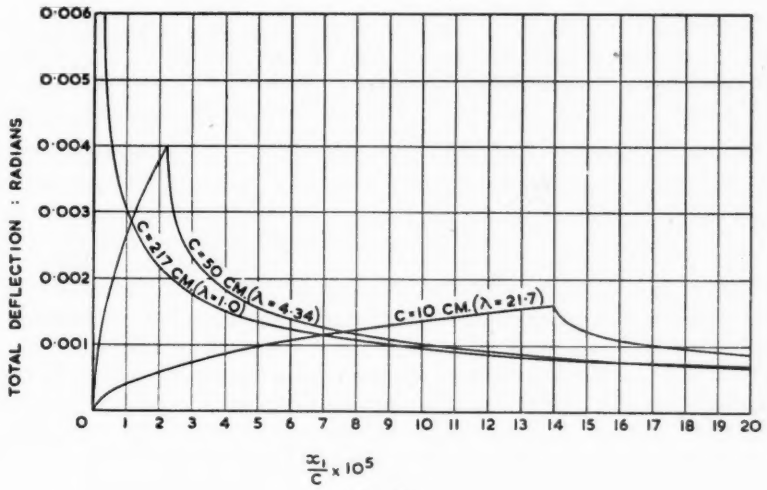

Fig. 20.
Refraction in a curved shock wave (convex on low-density side), $s=\Delta\rho/\rho=0.1$.

cm., or about twice the wavelength of light. Obviously, such a very narrow line will not be represented correctly by the optical system; it will appear as a much wider band in which there is a comparatively small change of illumination, instead of a very narrow band with a large change of illumination.

For a nearly two-dimensional shock wave the radius of curvature will be much greater. Referring to Fig. 15 it is seen that for $s=1$ and large radius the deflection exceeds 0.01 radian for a range of x_1/c between about 0.1×10^{-4} and 1.3×10^{-4}. If $c=100$ cm. this gives a band of width 0.12 mm. for which the deflection exceeds 0.01. This explains why nearly two-dimensional shock waves have a much greater effect than axially symmetrical waves, and often cause overloading of the schlieren system.

In considering the weakest shock wave that can be detected in axially symmetrical flow, the points of interest are the maximum deflection for a given shock wave and the value of x_1 at which this occurs. (This value of x_1 gives an indication of the width of the band for which the deflection is near the maximum value).

For a weak shock wave with c small, $\lambda >> 1$. Thus at the point of maximum deflection

$$\varepsilon_{max.}=\frac{2\alpha_1}{\lambda-1}\simeq\frac{2\alpha_1}{\lambda} \quad . \quad . \quad . \quad . \quad . \quad (37)$$

and

$$\alpha_1^2=2\theta\frac{\lambda-1}{\lambda}\simeq 2\theta.$$

Therefore

$$x_1\simeq t \quad . \quad . \quad . \quad . \quad . \quad . \quad (38)$$

Since the origin O is on the convex side of the shock wave (Fig. 18), the ray for which $x_1 \approx t$ just touches the concave side of the wave (if $\Delta\mu \longrightarrow 0$) and this condition gives maximum deflection. Also

$$\varepsilon_{max.} = \frac{2(2\theta)^{\frac{1}{2}}}{\lambda} = 2\left(\frac{2\Delta\mu}{\lambda}\right)^{\frac{1}{2}} \quad . \quad . \quad . \quad . \quad (39)$$

Therefore for a given strength of shock wave $\varepsilon_{max.} \propto \lambda^{-\frac{1}{2}} \propto c^{\frac{1}{2}}$.

Equation (38) shows that the width of the band in which the deflection is comparable with the maximum deflection is only of the order of the thickness of the shock wave. This is also true for a strictly plane shock wave, but it is probable that the shock waves in so-called two-dimensional flow are nearly always curved (with a large radius), so that the width of the band giving reasonably large deflections is considerably greater than the thickness of the shock wave. This explains why a weak shock wave of a given strength can be detected more easily in nearly two-dimensional flow than in axially symmetrical flow.

7. Boundary Layers

Boundary layers show clearly in schlieren photographs if the slits or knife-edges are arranged so that the density gradient in the boundary layer gives *more* light on the photograph. The light boundary layer then shows up clearly beside the dark shadow of the solid object. In a thin boundary layer (*e.g.* near the sharp leading edge of an aerofoil or wedge) the density gradient is very large. This may perhaps cause overloading of the schlieren system in some cases, but if this does occur it is not obvious, because its only effect is to cause a very slight apparent displacement of the dark boundary.

When a laminar boundary layer interacts with a shock wave, a dark line is sometimes seen near the middle of the boundary layer. This is shown, for example, in a paper by Liepmann[7]. The effect may possibly be caused by overloading of the schlieren system; the separating boundary layer may have a temperature distribution for which the gradient is a maximum near the middle. As in other cases, the possibility of overloading can be investigated by taking another photograph with the sign of the sensitivity reversed, making the boundary layer appear dark. If there is no light line in the dark boundary layer, corresponding to the dark line observed in the light boundary layer, there must have been overloading of the schlieren system.

Calculations of the deflection of a light ray by a boundary layer show that the deflection is large enough to be noticeable, with the sensitivity normally used in wind-tunnel work, in a region whose thickness is of the same order of magnitude as the total thickness of the boundary layer. Thus the apparent thickness of the boundary layer as seen on a schlieren photograph is a measure of the order of magnitude of the total boundary-layer thickness.

8. Dimensions of the Optical System

It is of interest to consider the effects of the dimensions of the system on the sensitivity and range, and in particular on the conditions for "overloading."

The effects of changing the dimensions of the light source and slits have already been discussed. Holder and North[3] have pointed out that a reduction of the focal length of the first lens or mirror increases the magnification of the image of the first slit, and hence increases the maximum range obtainable with a given source of light.

Now consider the effect of changing the diameter of the field of view (D). The relative aperture of the lenses or mirrors must be kept approximately constant, in order to avoid large aberrations while keeping the apparatus reasonably compact. Hence the length L in equation (3) will be roughly proportional to D. If the dimensions of the slits and light source are also made proportional to D, the range (expressed in terms of ε') will be constant. If D and L are *reduced*, keeping the dimensions of the slits and light source constant, the range will be *increased*.

In many cases the field of view is required to include the whole height (h) of the working section of a wind tunnel. Then, for geometrically similar wind tunnels

$$L \propto D \propto h \propto d,$$

where d is the width of the working section, measured in the direction of the light beam.

It has been shown that for a strong shock wave the deflection ε is independent of the width of the tunnel (d), unless d is very small. Hence for this case

$$\varepsilon L \propto h,$$

and if the dimensions of the slits and light source are made proportional to h, the range and sensitivity (and hence the conditions for overloading) are independent of size.

For a uniform density gradient of large extent, however (*e.g.* the outer part of a Prandtl-Meyer expansion)

$$\varepsilon \propto d \propto h$$

and hence

$$\varepsilon L \propto h^2.$$

Thus the ratio of the deflection at a strong shock wave (usually large) to the deflection in a uniform density gradient of large extent (usually small) *decreases as the size of the wind tunnel is increased.* This means that for nearly two-dimensional flow the problems connected with overloading of the schlieren system become less serious as the size of the wind tunnel is increased.

ACKNOWLEDGMENT

The photographs reproduced in Figs. 4 to 7 were taken by Mr. O. Bardsley.

REFERENCES

1. SCHARDIN, H. (1942). Die Schlierenverfahren und ihre Anwendungen. Ergebnisse der exakten Naturwissenschaften, Vol. 20, pp. 303-439, 1942.
(R.A.E. Library Translation No. 122, A.R.C. 10724.)

2. BARNES, N. F. and BELLINGER, S. L. (1945). Schlieren and Shadowgraph Equipment for Air Flow Analysis. *Journal of Optical Society of America*, Vol. 35, pp. 497-509, 1945.

3. HOLDER, D. W. and NORTH, R. J. (1950). The Toepler Schlieren Apparatus. A.R.C. 13068, 1950.

4. SPEAK, G. S. and WALTER, D. J. (1950). Optical Considerations and Limitations of the Schlieren Method. R.A.E. Technical Note. I.A.P. 968, A.R.C. 13066, 1950.

5. BARDSLEY, O. (1951). The Conditions at a Sharp Leading Edge in Supersonic Flow. *Philosophical Magazine*, Vol. 42, pp. 255-262, 1951.

6. TAYLOR, G. I. and MACCOLL, J. W. (1935). *Aerodynamic Theory* (edited by W. F. Durand), Berlin, 1935, Vol. III, Div. H., pp. 218-222.

7. LIEPMANN, H. W. (1950). Boundary-Layer Shock-Wave Interaction. U.S. Naval Ordnance Lab. Report No. 1133, 1950.

Note on Mr. Mair's Paper

D. W. HOLDER and R. J. NORTH

(*Aerodynamics Division, National Physical Laboratory*)

The scheme for taking a pair of photographs of opposite sensitivity for each flow is a useful one in certain problems, but is unsuitable for use in a routine experiment (in which several hundred photographs may be required) because of the additional labour and expense involved. This is particularly so because, apart from resolving any ambiguity which may occur in regions of intense density gradient, the second photograph usually will not give much additional information about the flow. Moreover, the scheme does not overcome the other difficulties associated with the use of an apparatus having inadequate range. The most serious of these is that the apparatus is unable to differentiate between the strengths of density gradients when they exceed the value which deflects the image of the light source completely on to, or away from, the knife edge. Photographs illustrating this effect are reproduced in Ref. 1; the use of inadequate range (*i.e.* too much sensitivity) is a common fault in wind-tunnel photographs.

The best procedure is to avoid the difficulty described by Mr. Mair by providing a working range which is large enough for no serious "overloading" to occur. In a Toepler apparatus fitted, as is now usual, with a knife edge (not a slit) in the focal plane of the second mirror, the working range is determined mainly (for a description of the influence of the other relevant dimensions *see* Ref. 1) by the diameter of the focusing lens and the height of the image of the light source. There is no serious difficulty in obtaining a sufficiently large lens particularly if, as is usually the case, it is of low power (and can thus consist of a single element), and it is felt that the provision of a lens of adequate diameter is essential in any schlieren apparatus which is to be used in routine work.

If the working range is made just large enough for no "overloading" to occur, the sensitivity of the apparatus will be fixed automatically, but in our experience it will usually remain adequate over the whole field. In the unusual case of additional sensitivity being required to observe the flow in part of the field, it is simple to take an additional photograph with higher sensitivity, and since the diameter of the focusing lens is adequate, no possibility of an ambiguity of the type described by Mr. Mair will arise. The most convenient way of altering the sensitivity and working range of a Toepler system is to adjust the height of the image of the source. A simple way of doing this is to rotate the source (or the slit in a condenser-slit arrangement as shown in Fig. 1 of Mr. Mair's paper) about the axis of the optical system. In this way the effective height of the image of the source can be

Note received August 1952.

[The Aeronautical Quarterly, Volume IV, August 1952]

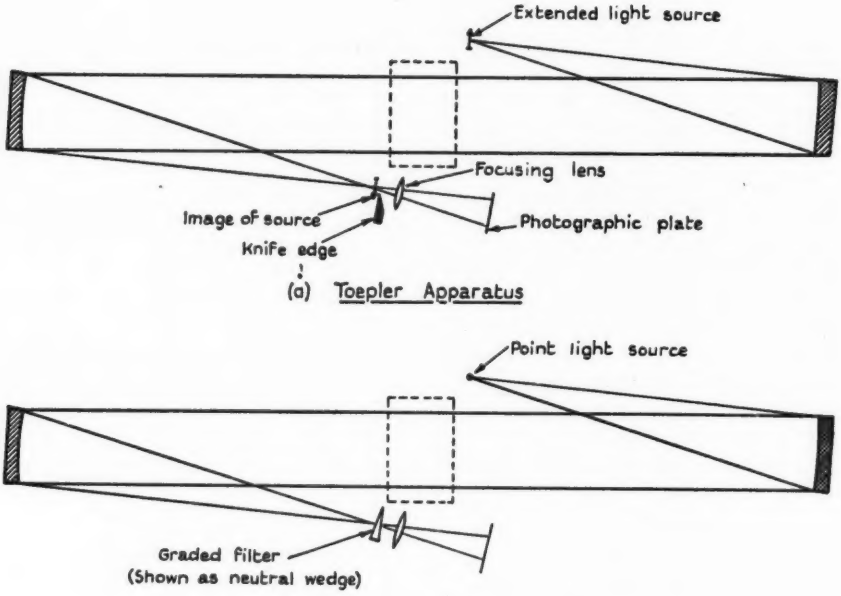

Fig. 1.

Diagrams of the schlieren systems.

adjusted to any value between its length and its breadth. Using this arrangement the greatest range obtainable in a given Toepler apparatus is limited by the length of the source image. This is an important matter in tunnel installations where the deflections of the light are large, because it is difficult to make this dimension very large without running into other difficulties.

An alternative schlieren system has recently been developed[2] at the National Physical Laboratory to avoid this difficulty. The disposition of the components (Fig. 1) is similar to that of a conventional Toepler apparatus except that a neutral filter screen is used instead of a knife edge in the focal plane of the second mirror. The optical density in the filter is graded so that the fraction of the incident light which is transmitted varies as the image of the light source moves across the filter. The image on the photographic plate thus resembles that obtained with a Toepler apparatus. There is no limitation to the size of the filter, and any required working range can be achieved by using a filter of suitable dimensions; with a given filter the sensitivity and working range can be adjusted within limits by tilting the filter about an axis perpendicular to the optical axis and to the optical density gradient in the filter. A point light source can be used, and this enables the same apparatus to

be used for direct shadow photography by placing the photographic plate immediately behind the working section of the wind tunnel.

In addition to overcoming the problem of providing adequate working range, the neutral filter apparatus gives photographs which appear considerably sharper than those obtained with a Toepler apparatus. The reason for this increase of resolving power is not yet clearly understood, but it is probably associated with the different effects of diffraction at a neutral screen and at a knife edge. If a linearly-graded filter is used an additional advantage of a practical nature is that the sensitivity will not change (as it does in a Toepler apparatus) because of a misalignment which displaces the image of the source relative to the filter. This feature may be valuable in certain wind-tunnel installations where there may be a considerable amount of vibration. Non-linear gradings may often be preferable because of other considerations, however, and this matter is discussed in Ref. 2. For example, it is possible to provide high sensitivity for the small density gradients in the flow while retaining adequate range for larger density gradients which may also be present.

The graded filters used so far have been made by printing through a neutral glass wedge on to photographic plates. This provides a convenient method for making large numbers of filters, and also enables the range and sensitivity to be altered by altering the photographic enlargement used in the printing process. A more flexible method, which would enable filters with a wide range of different characteristics to be produced, would be to use a cam-driven shutter moving across a photographic plate on the easel of an enlarger, and this and other methods are being investigated.

The graded-filter principle may also be applied to a colour-schlieren apparatus if a source of white light is used in the schlieren apparatus in conjunction with a filter graded in the colours of the spectrum instead of a neutral screen. The differing density gradients in the flow field are then represented in different colours on the viewing screen; this has been achieved previously[3] by means of a dispersing prism placed in front of the source of white light, and a slit instead of the knife edge of a Toepler apparatus in the focal plane of the second mirror. The filter screens for the colour apparatus have been made by photographing pure spectra of various sizes on 35 mm. colour film.

The work described forms part of the research programme of the Aerodynamics Division of the National Physical Laboratory, and this note is published with the permission of the Director.

REFERENCES

1. HOLDER, D. W. and NORTH, R. J. (1950). The Toepler Schlieren Apparatus. A.R.C. 13068, 1950. To be published in the R. & M. Series.
2. NORTH, R. J. (1952). Schlieren Systems Using Graded Filters. A.R.C. 15099, 1952.
3. HOLDER, D. W. and NORTH, R. J. (1950). A Colour Schlieren Apparatus. A.R.C. 12856, 1950. *See also Nature*, Vol. 169, p. 466, 1952. For colour prints *see The Aeroplane*, Vol. 82, No. 2111, 4th January 1952.

On the Conditions under which Energy can be Extracted from an Air Stream by an Oscillating Aerofoil

N. C. LAMBOURNE, B.Sc.

(Aerodynamics Division, National Physical Laboratory)

Summary: A simple system is considered in which a two-dimensional aerofoil in an air stream performs a coupled oscillation in pitching and vertical translation. An expression for the energy transferred from stream to aerofoil is derived as a function of the frequency parameter, amplitude ratio and phase difference, and the conditions under which positive energy can be extracted from the stream are investigated. The connection between the case of zero energy transfer, and the critical flutter state is noted.

Attention is confined to the incompressible case and diagrams show the results obtained when the air loads are obtained from (*i*) vortex sheet theory and (*ii*) the classical set of constant derivatives.

Vortex sheet theory leads to the conclusion that, provided that the amplitude ratio and phase difference satisfy certain conditions, energy can be extracted from the stream for all values of the frequency parameter. On the other hand, the use of a set of constant derivatives places an upper limit on the values of the frequency parameter for the extraction of energy.

1. Introduction

A consideration of the energy transfer that occurs when an aerofoil oscillates in an air stream is able to throw some light on the physics of the flutter phenomenon[1, 2, 3, 4], and the condition that the total energy transferred from stream to aerofoil per cycle of motion is zero is of particular interest, since it is satisfied at the critical flutter state. The conditions under which a two-dimensional aerofoil performing a coupled oscillation in pitching and vertical translation can absorb energy were considered by Küssner[2], but this work appears to have received only slight attention in Great Britain. Duncan[3] has described a simple "flutter engine" for the extraction of energy from an air stream and has referred to the manner in which the phase difference between the two motions influences the energy input. Frazer[5] has shown that the power required to maintain forced oscillations of an aerofoil in an air stream may be negative at speeds below the critical flutter speed.

Paper received July 1951.

[The Aeronautical Quarterly, Volume IV, August 1952]

The present paper follows a line suggested by Küssner's work, and gives the conditions under which it is possible to extract energy from the stream.

Notation

θ	angle of pitch (incidence) } *See* Fig. 1
z	vertical displacement of reference centre O } *See* Fig. 1
θ_0, z_0	amplitudes of θ and z
c	aerofoil chord
R	$z_0/(\theta_0 c)$, non-dimensional amplitude ratio
ρ	air density
V	air velocity
p	angular frequency
t	time
ε	the phase lead of z with respect to θ
X	$R\cos\varepsilon$
Y	$R\sin\varepsilon$
a, b	co-ordinates of the centre of a circle in the XY-plane
r	radius of a circle
ω	pc/V, non-dimensional frequency
M	aerodynamic moment per unit span } *See* Fig. 1
Z	aerodynamic normal force (lift) per unit span } *See* Fig. 1
$Z_z, Z_{dz/dt}, Z_{d^2z/dt^2}$ $Z_\theta, Z_{d\theta/dt}, Z_{d^2\theta/dt^2}$ $M_z, M_{dz/dt}, M_{d^2z/dt^2}$ $M_\theta, M_{d\theta/dt}, M_{d^2\theta/dt^2}$	derivative coefficients
$\overline{Z}_2, \overline{Z}_3, \overline{Z}_4$ $\overline{M}_1, \overline{M}_2, \overline{M}_4$	non-dimensional force and moment coefficients defined by equations (7)
$l_z', l'_{dz/dt}, l_\alpha', l'_{d\alpha/dt}$ $m_z', m'_{dz/dt}, m_\alpha', m'_{d\alpha/dt}$	non-dimensional derivative coefficients (*see* Section 4)
S, T_1, T_2, U	coefficients defined by equation (9)
A, B	functions of ω
W	work done by air forces per cycle per unit span
w	$W/(\pi^2\rho c^2 V^2\theta_0^2)$, non-dimensional form of W
k	a factor applied to direct dampings $Z_{dz/dt}$ and $M_{d\theta/dt}$

2. The Conditions for Energy Extraction

In the system under consideration an aerofoil in an air stream of velocity V is performing, by means of some arbitrary agency*, coupled simple harmonic motions in pitching and vertical translation (*see* Fig. 1). The flow is regarded as two-dimensional.

The motion of the aerofoil is defined by

$$\theta = \theta_0 \sin pt \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$$z = z_0 \sin (pt + \varepsilon) \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where θ is the angle of pitch and z the vertical displacement at some reference point O in the chord.

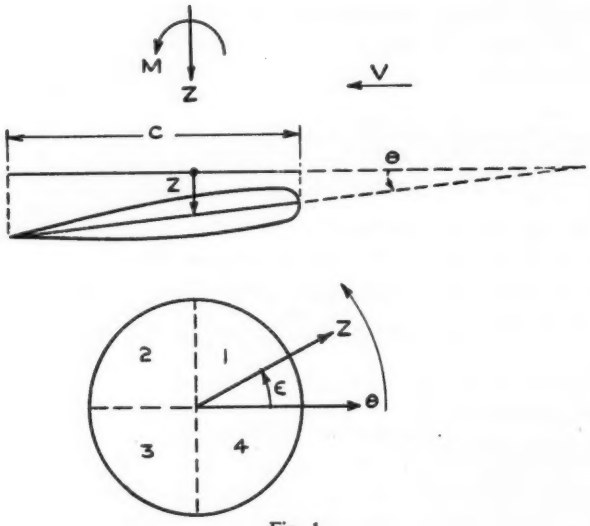

Fig. 1.
Definitions.

Then if the aerodynamic force and moment referred to point O are respectively Z and M per unit span, the work done by the air forces per cycle of motion, per unit span, is given by

$$W = \oint \left(Z\frac{dz}{dt} + M\frac{d\theta}{dt} \right) dt . \quad . \quad . \quad . \quad . \quad (3)$$

*For example, a two-dimensional version of the "flutter engine" referred to by Duncan[3].

This represents the energy that is extracted from the stream, and which in some manner is absorbed by the aerofoil or its activating agency.

The force and moment may be expressed in terms of derivative coefficients as follows:—

$$Z = Z_z z + Z_{dz/dt}\dot{z} + Z_{d^2z/dt^2}\ddot{z} + Z_\theta\theta + Z_{d\theta/dt}\dot{\theta} + Z_{d^2\theta/dt^2}\ddot{\theta} \quad . \quad . \quad (4)$$

$$M = M_z z + M_{dz/dt}\dot{z} + M_{d^2z/dt^2}\ddot{z} + M_\theta\theta + M_{d\theta/dt}\dot{\theta} + M_{d^2\theta/dt^2}\ddot{\theta} \quad . \quad . \quad (5)$$

By substitution in equation (3) and by performing the integration, the following expression for the energy is obtained:—

$$\frac{W}{\pi} = pZ_{dz/dt}z_0^2 + [p(Z_{d\theta/dt} + M_{dz/dt})\cos\varepsilon + \{(M_z - Z_\theta) - p^2(M_{d^2z/dt^2} - Z_{d^2\theta/dt^2})\}\sin\varepsilon]\, z_0\theta_0 + \\ + pM_{d\theta/dt}\theta_0^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

The non-dimensional forms of the derivative coefficients involved in equation (6) are, in the notation of Jones[6], as follows:—

$$\left.\begin{aligned}
\bar{Z}_2 &= -\frac{\omega Z_{dz/dt}}{\pi\rho c V} \\
\bar{Z}_3 &= -\frac{Z_\theta}{\pi\rho c V^2} + \frac{\omega^2 Z_{d^2\theta/dt^2}}{\pi\rho c^3} \\
\bar{Z}_4 &= -\frac{\omega Z_{d\theta/dt}}{\pi\rho c^2 V} \\
\bar{M}_1 &= -\frac{M_z}{\pi\rho c V^2} + \frac{\omega^2 M_{d^2z/dt^2}}{\pi\rho c^3} \\
\bar{M}_2 &= -\frac{\omega M_{dz/dt}}{\pi\rho c^2 V} \\
\bar{M}_4 &= -\frac{\omega M_{d\theta/dt}}{\pi\rho c^3 V}
\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

In general, on the basis of vortex sheet theory, the quantities on the left hand sides of equations (7) are functions of ω.

By substitution in equation (6) the following non-dimensional expression for the energy is obtained:—

$$w = \frac{W}{\pi^2\rho c^2 V^2\theta_0^2} = SR^2 + (T_1\cos\varepsilon + T_2\sin\varepsilon)R + U \quad . \quad . \quad . \quad (8)$$

where $R = \frac{z_0}{\theta_0 c}$ (the non-dimensional amplitude ratio)

$$\left.\begin{aligned} S &= -\bar{Z}_2 \\ T_1 &= -\bar{Z}_4 - \bar{M}_2 \\ T_2 &= +\bar{Z}_3 - \bar{M}_1 \\ U &= -\bar{M}_4 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

For energy to be extracted from the stream, the right hand side of equation (8) must be positive. It will be noticed that $-S$ and $-U$ correspond to the direct aerodynamic dampings, while T_1 corresponds to the cross dampings, and T_2 to the aerodynamic cross stiffnesses and inertias. It will be assumed that the direct dampings are positive so that both S and U will be negative*, and thus, since R is an amplitude ratio and therefore positive, a necessary, but not sufficient condition for the extraction of energy from the stream is

$$T_1 \cos \varepsilon + T_2 \sin \varepsilon > 0 . \quad . \quad . \quad . \quad (10)$$

The energy transfer changes sign when $w = 0$ or

$$SR^2 + (T_1 \cos \varepsilon + T_2 \sin \varepsilon) R + U = 0 . \quad . \quad . \quad . \quad (11)$$

Since R must be real and positive, when $w = 0$ the following condition holds

$$T_1 \cos \varepsilon + T_2 \sin \varepsilon > 2\sqrt{(SU)} . \quad . \quad . \quad . \quad (12)$$

By inserting the appropriate values of T, S and U in the inequality (12) it is possible to find, for any value of ω, the range of phase angles over which the condition $w = 0$ may be satisfied.

The optimum phase angle for the extraction of energy is obtained

from $$\frac{\partial w}{\partial \varepsilon} = 0$$

$$\frac{\partial^2 w}{\partial \varepsilon^2} < 0 ,$$

and is given by $$\tan \varepsilon_{opt} = \frac{T_2}{T_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

*This is true except for very low values of ω, when vortex sheet theory indicates that U may be positive for some values of the reference point.

and in the usual notation the quadrant within which ε_{opt} lies is given by the following table

T_1	T_2	*Quadrant*
+ve	+ve	1st
−ve	+ve	2nd
−ve	−ve	3rd
+ve	−ve	4th

It will be noted in passing that the optimum phase angle for energy transfer *from* aerofoil *to* stream is

$$\varepsilon'_{opt}=\varepsilon_{opt}+\pi.$$

If the optimum value for energy extraction from the stream is assigned to the phase angle (*i.e.* $\varepsilon_{opt}=\tan^{-1} T_2/T_1$) the energy expression given by equation (8) becomes

$$w_{\varepsilon opt}=SR^2+(T_1^2+T_2^2)^{\frac{1}{2}}R+U \quad . \quad . \quad . \quad . \quad (14)$$

The limiting values of R for which it is possible to satisfy the condition $w=0$ are then given by

$$SR^2+(T_1^2+T_2^2)^{\frac{1}{2}}R+U=0 \quad . \quad . \quad . \quad . \quad (15)$$

Also when

$$R=-\frac{(T_1^2+T_2^2)^{\frac{1}{2}}}{2S} \quad . \quad . \quad . \quad . \quad . \quad (16)$$

the expression for $w_{\varepsilon opt}$ shows a maximum which is given by

$$w_{max}=-\frac{T_1^2+T_2^2}{4S}+U \quad . \quad . \quad . \quad . \quad . \quad (17)$$

3. Results Based on Vortex Sheet Theory

The reference point was chosen at $0.3c$ from the leading edge and values of S, T_1, T_2 and U were obtained from the tables of air load coefficients given by Jones[6]. For any particular value of ω the relation between the phase and amplitude ratio when $w=0$ may be shown graphically in the following manner. By writing

$$X=R\cos\varepsilon$$

$$Y=R\sin\varepsilon$$

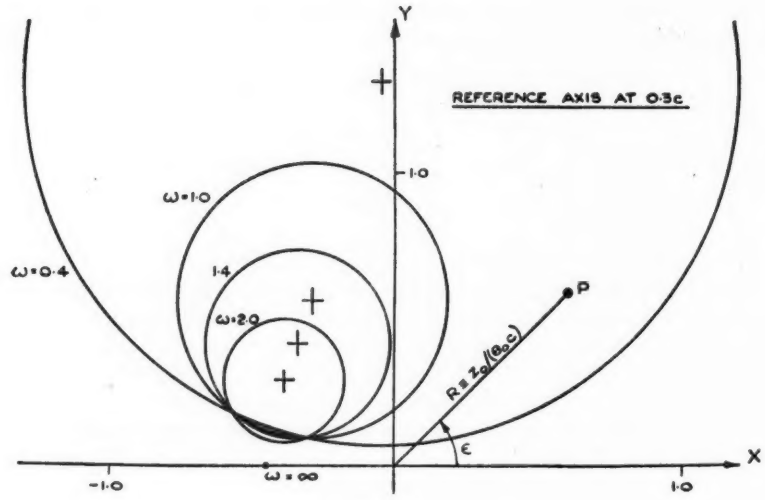

Fig. 2.

Relations between amplitude ratio R and phase ϵ for condition $w=0$ based on vortex sheet theory. (Energy is extracted from the stream when P lies within the appropriate circle.)

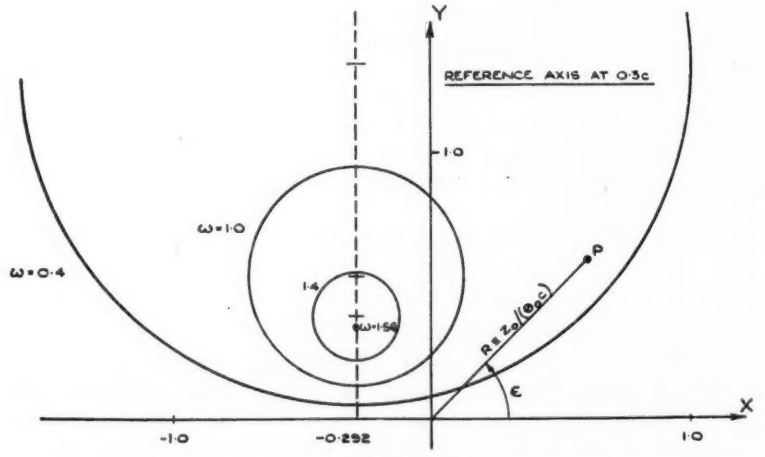

Fig. 3.

Relations between amplitude ratio R and phase ϵ for condition $w=0$ based on the "classical" constant derivatives. (Energy is extracted from the stream when P lies within the appropriate circle.)

equation (11) becomes the circle

$$(X-a)^2+(Y-b)^2=r^2 \qquad . \quad . \quad . \quad (18)$$

where the co-ordinates of the centre and the radius are given by

$$a=-\frac{T_1}{2S} \qquad =-\frac{\overline{M}_2+\overline{Z}_4}{2\overline{Z}_2}$$

$$b=-\frac{T_2}{2S} \qquad =-\frac{\overline{M}_1-\overline{Z}_3}{2\overline{Z}_2}$$

$$r^2=\frac{T_1^2+T_2^2}{4S^2}-\frac{U}{S}=\frac{(\overline{M}_2+\overline{Z}_4)^2+(\overline{M}_1-\overline{Z}_3)^2}{4\overline{Z}_2^2}-\frac{\overline{M}_4}{\overline{Z}_2}.$$

A condition under which energy is extracted from the stream is then represented graphically by a point (polar co-ordinates $R\varepsilon$) lying within the circle. The circles for a few values of ω are shown in Fig. 2.

Figures 4 and 5 show respectively the limiting phase angles and the limiting amplitude ratios for which it is possible to satisfy the condition $w=0$. Further curves are also included in these figures to show the corresponding ranges when the direct dampings $Z_{dz/dt}$ and $M_{d\theta/dt}$ are increased by a factor k. These additional curves will be referred to later in Section 5.

Küssner considered a relation similar to that of equation (15) and he gave the numerical results as a curve showing amplitude ratio plotted against $2/\omega$. When the "vortex sheet" results given in Fig. 5 are replotted in this form, they are identical with those given by Küssner.

The case $\omega \longrightarrow \infty$ is most easily dealt with by considering the reference point at the mid-chord position. Then, in the notation of Jones[6], the coefficients of equation (8) are

$$S=-Z_2 \qquad =-\omega A$$

$$T_1=-Z_4-M_2=-\frac{\omega}{4}+B$$

$$T_2=+Z_3-M_1=A+\frac{\omega B}{2}$$

$$U=-M_4 \qquad =\frac{\omega A}{16}-\frac{\omega}{16}-\frac{\omega D}{4}$$

where $A \longrightarrow 0.5$ and $B \longrightarrow 0$ when $\omega \longrightarrow \infty$.

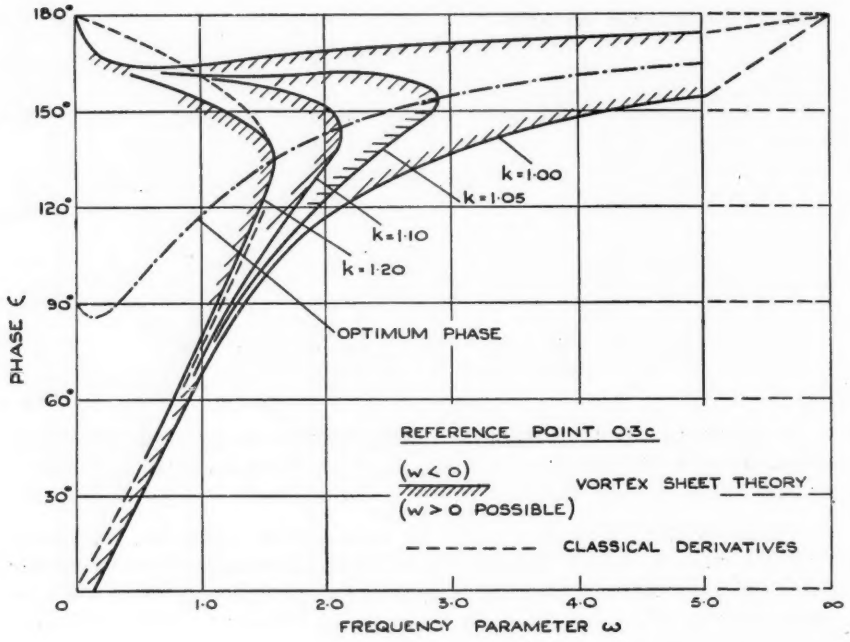

Fig. 4.

Phase angle ranges over which energy may be extracted from the stream.
Direct dampings increased by factor k.

Equation (8) may then be rewritten as

$$w=-\omega\left\{AR^2+\left[\left(\frac{1}{4}-\frac{B}{\omega}\right)\cos\varepsilon-\left(\frac{A}{\omega}+\frac{B}{2}\right)\sin\varepsilon\right]R-\frac{A}{16}+\frac{1}{16}+\frac{B}{4}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

The condition that the left hand side of this equation is zero when $\omega\longrightarrow\infty$ is, by insertion of the limits of A and B,

$$16R^2+(8\cos\varepsilon)R+1=0\,. \quad . \quad . \quad . \quad (20)$$

The only value of ε that will admit real and positive roots for R is π, and this leads to a pair of equal roots

$$R=\frac{1}{4}\,.$$

Thus when $\omega \longrightarrow \infty$, the condition $w=0$ corresponds to a pure pitching oscillation of the aerofoil about an axis at $0.75c$ from the leading edge. This result was referred to by Küssner[2] and has been deduced by a different method by Biot and Arnold[7].

The values of R and ε appropriate to a reference point at $0.3c$ are obtained directly from the foregoing result and are 0.45 and π respectively.

The results based on vortex sheet theory derivatives thus suggest that, provided certain conditions regarding phase and amplitude are fulfilled, energy may be extracted from the stream for any value of ω.

4. Results Based on a Set of Constant Derivative Coefficients

The constant non-dimensional derivative coefficients are defined as follows:—

$$l_z' = -\frac{Z_z}{\rho V^2}, \qquad m_z' = \frac{M_z}{\rho V^2 c}$$

$$l'_{dz/dt} = -\frac{Z_{dz/dt}}{\rho V c}, \qquad m'_{dz/dt} = \frac{M_{dz/dt}}{\rho V c^2}$$

$$l_\alpha' = -\frac{Z_\theta}{\rho V^2 c}, \qquad m_\alpha' = \frac{M_\theta}{\rho V^2 c^2}$$

$$l'_{d\alpha/dt} = -\frac{Z_{d\theta/dt}}{\rho V c^2}, \qquad m'_{d\alpha/dt} = \frac{M_{d\theta/dt}}{\rho V c^3}$$

$$Z_{d^2z/dt^2} = Z_{d^2\theta/dt^2} = M_{d^2z/dt^2} = M_{d^2\theta/dt^2} = 0.$$

The expression for W, the work done by the air forces per cycle of motion per unit span, is given by

$$w = \frac{W}{\pi^2 \rho c^2 V^2 \theta_0^2} = -\frac{1}{\pi} \{ \omega l'_{dz/dt} R^2 + [\omega (l'_{d\alpha/dt} - m'_{dz/dt}) \cos \varepsilon - (l_\alpha' + m_z') \sin \varepsilon] R - \omega m'_{d\alpha/dt} \} \quad . \quad . \quad . \quad . \quad . \quad (21)$$

The optimum phase difference for extraction of energy from the stream is given by

$$\tan \varepsilon_{opt} = -\frac{(l_\alpha' + m_z')}{\omega (l'_{d\alpha/dt} - m'_{dz/dt})} \quad . \quad . \quad . \quad . \quad . \quad (22)$$

and the energy extracted under this condition is given by

$$w_{opt} = -\frac{1}{\pi} \{ \omega l'_{dz/dt} R^2 - [\omega^2 (l'_{d\alpha/dt} - m'_{dz/dt})^2 + (l_\alpha' + m_z')^2]^{\frac{1}{2}} R - \omega m'_{d\alpha/dt} \} \quad . \quad (23)$$

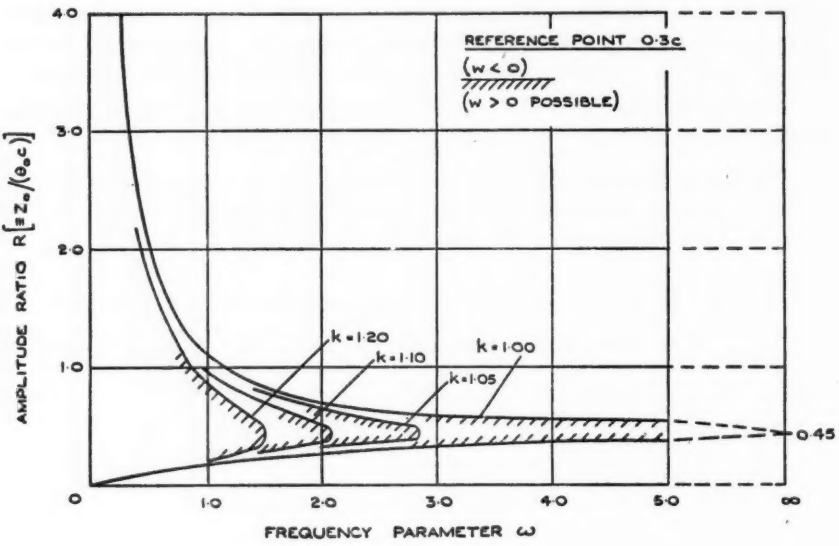

Fig. 5.

Amplitude ratios at which energy may be extracted from the stream. Based on "vortex sheet theory" and with direct dampings increased by factor k.

If $l'_{dz/dt}$ is positive and $m'_{d\alpha/dt}$ is negative then, since R must be real and positive, a condition that the right hand side of equation (23) is positive is

$$\omega^2 (l'_{d\alpha/dt} - m'_{dz/dt})^2 + (l_\alpha' + m_z')^2 > -4\omega^2 l'_{dz/dt} m'_{d\alpha/dt}.$$

Thus no energy can be extracted from the stream unless

$$(l_\alpha' + m_z')^2 + \omega^2 \{ (l'_{d\alpha/dt} - m'_{dz/dt})^2 + 4l'_{dz/dt} m'_{d\alpha/dt} \} > 0. \quad . \quad . \quad (24)$$

The relevant derivative coefficients of the classical set, first given by Duncan and Collar[8], refer to the leading edge and are as follows:—

$$l_{dz/dt} = 1.5, \qquad m_{dz/dt} = -0.375,$$
$$l_{d\alpha/dt} = 1.4, \qquad m_{d\alpha/dt} = -0.7,$$
$$l_\alpha = 1.6, \qquad m_z = 0.$$

By substitution in inequality (24) the following criterion for the extraction of energy is obtained:—

$$\omega < 1.56.$$

The phase range corresponding to a reference point at $0.3c$ for the condition $w=0$ has been obtained by a consideration of equation (18) and is shown in Fig. 4. As in the case of the "vortex sheet" results the relation between the phase and amplitude ratio for the condition $w=0$ may be shown graphically as a circle for each ω value on a polar ($R\varepsilon$) diagram.

The Cartesian co-ordinates of the centre are

$$-\frac{l'_{d\alpha/dt}-m'_{dz/dt}}{2l'_{dz/dt}}, \quad +\frac{l_\alpha'+m'_z}{2\omega l'_{dz/dt}}$$

and the radius is

$$\left\{\frac{(l'_{d\alpha/dt}-m'_{dz/dt})^2}{4l'^2_{dz/dt}}+\frac{(l_\alpha'+m_z')^2}{4\omega^2 l'^2_{dz/dt}}+\frac{m'_{d\alpha/dt}}{l'_{dz/dt}}\right\}^{\frac{1}{2}}.$$

The circles for a few values of ω are shown in Fig. 3.

Various sets of constant derivative coefficients have been quoted by Williams[9] and the maximum values of ω for energy extraction appropriate to these sets are given in Table I.

5. Critical Flutter State

If the aerofoil is now regarded as being divorced from the activating agency, but as part of an elastic-inertia system, which for the moment will be assumed free from structural damping, then the condition $w=0$ holds at the critical flutter state*. The quantities

frequency parameter	ω
amplitude ratio	R
phase lead	ε

are then related by equation (11) and, provided that two are known, the third quantity could be determined from a set of curves prepared on the lines of either Fig. 2 or Fig. 3.

For a practical system other dampings (*e.g.* structural damping) would be present in addition to the aerodynamic damping already considered, and at the critical flutter state a finite amount of energy would have to be extracted from the stream to be dissipated by the additional damping. The effect of this additional damping may be shown in general terms by considering what happens when the direct aerodynamic dampings $Z_{dz/dt}$ and $M_{d\theta/dt}$ are numerically increased while the other derivative coefficients remain unaltered.

*The state at which the amplitude of the flutter is just maintained.

TABLE I

MAXIMUM VALUES OF ω FOR EXTRACTION OF ENERGY FOR SETS OF VALUES OF THE LOCAL LEADING EDGE COEFFICIENTS. (*See* Ref. 9)

Origin (*See Ref. 9 for details*)	$l'_{dz/dt}$	$l'_{d\alpha/dt}$	l_α'	$-m'_{dz/dt}$	$-m'_{d\alpha/dt}$	$-m_\alpha'$	l_z', m_z'	*Maximum ω for energy extraction*
R. & M. 1155 (Duncan, R. & M. 1782)	1.5	1.4	1.6	0.37_5	0.70	0.40	↑	1.56
R. & M. 1155 ($f=F=\eta^{7/6}$)	1.1_0	1.0_5	1.3_3	0.34	0.54	0.26		2.00
Cicala (Experimental-theoretical mean)	1.7_3	1.6_2	1.7_3	0.43	0.71	0.43		2.05
Two dimensional vortex sheet theory							zero	
ω=1.0	1.8_8	1.7_2	2.0_9	0.47	0.82	0.52		1.79
ω=2.0	1.7_0	1.9_0	1.9_7	0.42_5	0.87	0.49		2.76
Jones, theoretical ($A=6$, $f=F=1$, $\gamma_t=1$)	1.5_5	1.6	1.8	0.38	0.74	0.45		2.20
"Mean experimental" (from derivative experiments)	1.6	1.6	1.9	0.40	0.72	0.47_5		2.44
"Proposed set" (from derivative and flutter experiments)	1.6	1.6	1.8	0.40	0.70	0.45	↓	2.60

Calculations of the phase and amplitude ratio ranges were made with the direct dampings increased in turn by the factors

$$k = 1.05,\ 1.10,\ 1.20,$$

and the results are included in Figs. 4 and 5. It is seen that for a system containing additional damping there is an upper limit to the frequency parameter range over which an instability can occur.

6. Concluding Remarks

1. The numerical results based on vortex sheet derivatives show that it is possible for a two-dimensional aerofoil performing coupled vertical translation and pitching oscillations to extract energy from an air stream at any air speed provided that the amplitude ratio and phase difference satisfy certain conditions. This is in agreement with the work of Frazer[5] which has shown that a practical aerofoil system involving stiffnesses and inertias can, for some conditions of excitation, extract energy at air speeds below its critical flutter speed.

2. Both the phase difference and amplitude ratio ranges for extraction of energy are narrowed down with increase of ω.

3. The result obtained when $\omega \longrightarrow \infty$ is interesting since it shows that as $V \longrightarrow 0$ the type of oscillation which allows energy to be extracted approaches pure pitching about an axis $0.75c$ from the leading edge. From this conclusion it may be inferred, as Küssner[2] pointed out, that a wing which has a nodal line at this position during a still air resonance is susceptible to flutter.

4. A point is now noted in connection with the flutter problem of a wing, any chordwise section of which has freedoms in

(*i*) vertical translation (due possibly to flexure, bodily roll and bodily translation)

(*ii*) pitching (due to torsion and bodily pitching).

At the critical flutter state there will, in general, be a positive transfer of energy from stream to wing for some chordwise section, and a negative transfer for others. On the usual strip theory basis, it is then true that for one chordwise section at least the appropriate values of R and ε must lie within the circle (*see* Fig. 2) corresponding to the current value of ω for the section. Thus a set of circles prepared on the lines of Fig. 2 could provide a check on the results of a flutter calculation that has determined the flutter mode as well as the critical speed and frequency.

5. On the basis of the classical set of constant derivatives energy cannot be extracted from the stream by a two-dimensional aerofoil if $\omega > 1.56$. Thus a flutter

calculation for the system using these derivatives could never yield a critical frequency parameter ω_c greater than this value. In the case of a tapered wing, treated on a strip theory basis, the limitation applies to critical frequency parameter appropriate to the smallest chord.

6. If a two-dimensional aerofoil-inertia-elastic system contains additional damping, there is, even on vortex sheet theory, an upper limit to the frequency parameter for which a maintained oscillation can occur.

ACKNOWLEDGMENTS

The author acknowledges the assistance of Mr. P. S. Pusey who carried out some of the computational work.

The work described was done in the Aerodynamics Division of the National Physical Laboratory, and it is published on the recommendation of the Aeronautical Research Council and by permission of the Director of the Laboratory.

REFERENCES

1. LOCKSPEISER, B. (1933). A Simple Approach to the Wing Flutter Problem. *Journal of the Royal Aeronautical Society*, July 1933.

2. KUSSNER, H. G. (1935). The Present Stage of Development of the Problem of Wing Flutter. Luftfahrtforschung No. 6, October 1935.

3. DUNCAN, W. J. (1944). The Fundamentals of Flutter. A.R.C. Report 7716, 1944.

4. GREIDANUS, J. H. (1940). De Arbeid, die de Luchtkrachten per Tijdseenheid op een Trillenden Vleugel Verrichten en de Kritische Trillingsvormen van een Vlengel. Report V, 1237. National Aeronautical Research Institute, Amsterdam, 1940.

5. FRAZER, R. A. (1939). On the Power Input required to Maintain Forced Oscillations of an Aeroplane Wing in Flight. R. & M. 1872, 1939.

6. JONES, W. P. (1942). Summary of Formulae and Notations used in Two Dimensional Derivative Theory. R. & M. 1958, 1942.

7. BIOT, M. A. and ARNOLD, LEE (1948). Low-speed Flutter and its Physical Interpretation. *Journal of the Aeronautical Sciences*, Vol. 15, No. 4, April 1948.

8. DUNCAN, W. J. and COLLAR, A. R. (1932). Present Position of Investigation of Airscrew Flutter. R. & M. 1518, 1932.

9. WILLIAMS, J. (1943). An Examination of Experimental Data relating to Flexural-torsional Wing Derivatives. R. & M. 1944, 1943.

Aerofoil Theory for Swallow Tail Wings of Small Aspect Ratio*

A. ROBINSON, Ph.D.

(*University of Toronto*)

Summary: A method is developed for the calculation of the aerodynamic forces acting on a "swallow tail" wing (Figs. 1(*a*) and 1(*b*)) of small aspect ratio. For a given incidence, lift and induced drag coefficient are, within the limits of the theory, proportional to aspect ratio and independent of Mach number. The exact solution is derived for the limiting case of small sweepback of the trailing edge. A numerical method is indicated for the more general case.

1. Introduction

In a remarkable paper published some years ago, R. T. Jones[1] put forward a theory for low aspect ratio pointed wings at all speeds, both below and above the speed of sound. The theory was based on the idea, due to Munk[2], that the induced velocity field set up by a slender moving body in a transverse plane, is essentially two-dimensional. In spite of the simplification involved in this idea, the results of Ref. 1 were borne out in a striking manner, both by relatively more exact theories and by experiment. The theory was extended by H. S. Ribner[3] to permit the calculation of the stability derivatives of low aspect ratio wings. It has also been used for the calculation of the effect of controls of different types by J. De Young[4] and M. D. Hodges[5]. G. N. Ward[6] has applied an equivalent method to problems of wing body interference at supersonic speeds in a paper which also includes a rigorous justification of the basic assumption already mentioned for the case under consideration.

The present paper concerns the calculation of the aerodynamic forces acting on small aspect ratio "swallow tail" wings such as depicted in Figs. 1(*a*) and 1(*b*). It is assumed that the outline of the wing plan form varies monotonically from a pointed nose to pointed tips (*i.e.* $dy/dx < 0$ along ABC in Fig. 1(*a*) and $dy/dx > 0$ along AFE). This case is outside the scope of the methods which are given in the papers mentioned, although for one particular plan form a solution for a mathematically equivalent problem is described in Ref. 7.

An exact solution will be derived for the limiting case of small sweepback of the trailing edge of a simple swallow tail wing (Fig. 1(*b*)).

---

*Modified version of College of Aeronautics Report No. 41, October 1950. Submitted for publication November 1950. The author is indebted to Mr. G. N. Ward and to members of the Royal Aircraft Establishment for critical comments.

Notation

x, y, z	rectangular axes; x parallel to the direction of flow, y to starboard, z upwards
Φ	total velocity potential
ϕ	induced velocity potential
V	free stream velocity
u, v, w	induced velocity components ($=\partial\phi/\partial x, \partial\phi/\partial y, \partial\phi/\partial z$ respectively)
Δp	pressure difference at a point of the aerofoil
ρ	air density
α	angle of incidence, positive in the nose-up sense
$a(x), b(x)$	spanwise co-ordinates of leading and trailing edges, respectively
c	local chord
c_0	root chord
η	$y+iz$
$\mathscr{R}, \mathscr{I}$	real and imaginary parts of
t	variable of integration in equation (4)
$A(t)$	function introduced in equation (4)
$g(t, \eta)$	function defined by equation (5)
k	$b(t)/a(t)$
K', E'	complementary complete elliptic integrals of modulus k
s	variable of integration in equation (5)
k'	$\sqrt{(1-k^2)}$
τ	variable of integration in equation (6)
τ'	τ/i
$F(\beta,k), E(\beta,k)$	elliptic integrals of the first and second kind
$f(\eta)$	an analytic function
$f'(\eta)$	$\partial f(\eta)/\partial\eta$
$a'(t)$	$\partial a(t)/\partial t$
a_0	$a(c_0)$
r	$a(t)/s$
$N(t, x)$	*see* equation (21)
a, b	$a(x), b(x)$ respectively, in Section 3
$l(x)$	lift per unit chord
$M(t, x)$	$N(t, x)/(K'-E')$
x', x''	particular values of x
ξ	intermediate value between x' and x''
γ	semi apex angle of wing
μ	$(c-c_0)/c_0$
d	span$=2(1+\mu)a_0$
S	wing area
σ	variable of integration, defined by $t=(1+\mu\sigma)c_0$
$M^*(\sigma, s)$	$a_0 M(t, x)$
$l^*(\sigma)$	$l(t)/\{2\pi\rho V^2 z a_0 \tan\gamma\}$

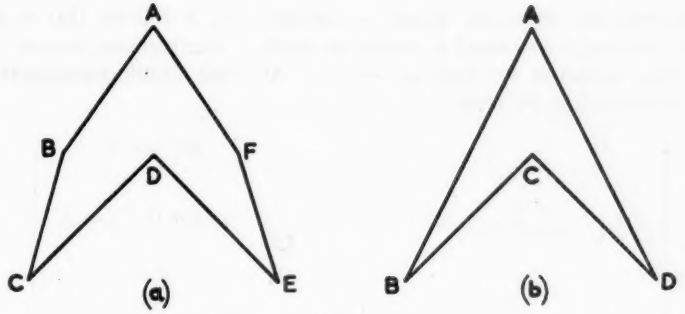

Fig. 1.

2. Analysis

The origin is taken at the apex of the wing, with the x-axis parallel to the direction of flow, the y-axis pointing to starboard, and the z-axis pointing upwards, so that the co-ordinate system is right handed. Let Φ be the total velocity potential and ϕ the induced velocity potential so that $\Phi = Vx + \phi$. The induced velocity components are denoted by $u = \partial\phi/\partial x$, $v = \partial\phi/\partial y$, $w = \partial\phi/\partial z$, so that the total longitudinal velocity equals $V + u$. By Bernoulli's equation, the pressure difference at a point of the aerofoil, Δp, is related to the longitudinal induced velocity by

$$\Delta p = -2\rho V u = -2\rho V \frac{\partial \phi}{\partial x} \quad . \quad . \quad . \quad . \quad (1)$$

In accordance with the introduction, the partial differential equation for ϕ is taken as

$$\frac{\partial^2 \phi}{\partial y^2} + \frac{\partial^2 \phi}{\partial z^2} = 0 . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The boundary condition at the aerofoil is

$$w = \frac{\partial \phi}{\partial z} = -V\alpha \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where α is the local incidence. Only the case of a flat wing, α=constant will be considered. Writing $y = \pm a(x)$, $0 \leqslant x \leqslant c$, and $y = \pm b(x)$, $c_0 \leqslant x \leqslant c$ for the spanwise co-ordinates of leading and trailing edges respectively, $a(0) = b(c_0) = 0$, and $a(c) = b(c)$. The complex variable $\eta = y + iz$ is also introduced so that ϕ may be regarded as a function of the two variables x and η, and may be written as the real part of a complex function of x and η which is an analytic function of η. ϕ must satisfy equations (2) and (3) and in addition must be such that Δp and hence $\partial\phi/\partial x$ vanish at the trailing edges of the aerofoil (Joukowski condition), while these quantities may become infinite at the leading edge. Finally, since the aerofoil does

not penetrate any transverse plane $x=\text{constant}<0$, it follows that ϕ must be constant, and may be assumed to vanish for such x. Furthermore, for any given x, ϕ must then vanish in the limit as $\eta\longrightarrow\infty$. All these conditions suggest that ϕ can be represented in the form

$$\phi(x,\eta)=\left\{\begin{array}{ll} 0 & \text{for } x<0 \\ \mathscr{R}\displaystyle\int_0^x \frac{A(t)}{\sqrt{\{[a(t)]^2-\eta^2\}}}\,dt & \text{for } 0<x<c_0 \\ \mathscr{R}\Big[\displaystyle\int_0^{c_0}\frac{A(t)}{\sqrt{\{[a(t)]^2-\eta^2\}}}\,dt+\int_{c_0}^{x}A(t)\,g(t,\eta)\,dt\Big] & \text{for } c_0<x<c\,, \end{array}\right\} \quad (4)$$

where A is a real function of its argument, which remains to be determined, and

$$g(t,\eta)=\int_0^{\eta}\left(K'-E'-\frac{\{[a(t)]^2-[b(t)]^2\}K'}{[a(t)]^2-s^2}\right)\frac{ds}{\sqrt{\{([a(t)]^2-s^2)(s^2-[b(t)]^2)\}}} \quad (5)$$

where K' and E' are the complementary complete elliptic integrals of modulus $k=b(t)/a(t)$. The sign of the integrand at the lower limit of integration ($s=0$) is made determinate by choosing $\sqrt{\{[a(t)]^2(-[b(t)]^2)\}}=i|ab|$ for that point. Substituting $s=b\,\mathrm{sn}(\tau,k)$ in (5) gives

$$\begin{aligned} g(t,\eta)&=\frac{1}{ia}\int_0^{\mathrm{sn}^{-1}\frac{\eta}{b}}\left(K'-E'-\frac{(1-k^2)K'}{1-k^2\,\mathrm{sn}^2(\tau,k)}\right)d\tau \\ &=\frac{1}{ia}\int_0^{\mathrm{sn}^{-1}\frac{\eta}{b}}\left(K'-E'-\frac{k'^2K'}{\mathrm{dn}^2(\tau,k)}\right)d\tau \\ &=\frac{1}{ia}\Big[(K'-E')\tau-K'\{E(\tau,k)-k^2\,\mathrm{sn}(\tau,k)\,\mathrm{cd}(\tau,k)\}\Big]_0^{\mathrm{sn}^{-1}\frac{\eta}{b}} \end{aligned} \quad (6)$$

where $k'^2=1-k^2$. Now if the path of integration is taken from 0 to $i\infty$ along the positive imaginary axis, the corresponding path in the τ-plane is the segment $(0,iK')$ on the positive imaginary axis. Putting $\tau=i\tau'$, then by Jacobi's transformation,

$$\begin{aligned} E(\tau,k)&=\tau+i\,\mathrm{dn}(\tau',k')\,\mathrm{sc}(\tau',k')-iE(\tau',k') \\ \mathrm{sn}(\tau,k)&=i\,\mathrm{sc}(\tau',k') \\ \mathrm{cn}(\tau,k)&=\mathrm{nc}(\tau',k') \\ \mathrm{dn}(\tau,k)&=\mathrm{dc}(\tau',k') \end{aligned}$$

so that

$$E(\tau,k)-k^2\,\mathrm{sn}(\tau,k)\,\mathrm{cd}(\tau,k)$$

$$=\tau-iE(\tau',k')+i\,[\mathrm{dn}(\tau',k')\,\mathrm{sc}(\tau',k')-k^2\,\mathrm{nd}(\tau',k')\,\mathrm{sc}(\tau',k')]$$

$$=\tau-iE(\tau',k')+i\,\mathrm{nd}(\tau',k')\,\mathrm{sc}(\tau',k')[\mathrm{dn}^2(\tau',k')-k^2]$$

$$=\tau-iE(\tau',k')+ik'^2\,\mathrm{sd}(\tau',k')\,\mathrm{cn}(\tau',k').$$

This expression tends to $iK'-iE'$ as τ tends to iK', since $\mathrm{cn}(K',k')=0$. Hence $g(t,\eta)\longrightarrow 0$ as $\eta\longrightarrow i\infty$.

Thus, by (5) and (6), $g(t,\eta)$ is one-valued, and is regular and vanishes at $\eta=\infty$, and it follows that the same applies to $\phi(x,\eta)$ as given by (4).

The longitudinal induced velocity at the aerofoil ($\eta=y$) is given by

$$u=\frac{\partial\phi}{\partial x}=\left\{\begin{array}{ll}\dfrac{A(x)}{\sqrt{\{[a(x)]^2-y^2\}}} & \text{for } 0<x<c_0\\[2ex] \mathscr{R}\,A(x)\,g(x,y) & \text{for } c_0<x<c\end{array}\right\} \quad . \quad . \quad (7)$$

Now the integrand of $g(t,\eta)$ is imaginary in the interval $-b<\eta<b$, and the integral is finite though improper at the trailing edges $\eta=\pm b$. It follows that the Joukowski condition is satisfied.

To find the normal induced velocity, (4) must be differentiated with respect to z. For any analytic function $f(\eta)$

$$\frac{\partial}{\partial z}\mathscr{R}f(\eta)=\mathscr{R}\frac{\partial f(\eta)}{\partial z}=\mathscr{R}\left[i\frac{\partial f(\eta)}{\partial\eta}\right]=-\mathscr{I}f'(\eta)$$

where $\mathscr{I}$ denotes the (real) coefficient of the imaginary part of a complex number, as usual. Hence

$$\frac{\partial\phi}{\partial z}=\left\{\begin{array}{ll} 0 & \text{for } x<0\\[2ex] -\mathscr{I}\left[\dfrac{\partial}{\partial\eta}\displaystyle\int_0^x\frac{A(t)}{\sqrt{\{[a(t)]^2-\eta^2\}}}\,dt\right], & \text{for } 0<x<c_0\\[2ex] -\mathscr{I}\left[\dfrac{\partial}{\partial\eta}\displaystyle\int_0^{c_0}\frac{A(t)}{\sqrt{\{[a(t)]^2-\eta^2\}}}\,dt+\dfrac{\partial}{\partial\eta}\displaystyle\int_{c_0}^x A(t)\,g(t,\eta)\,dt\right] & \text{for } c_0<x<c.\end{array}\right\} \quad (8)$$

$A(t)$ must now be determined from the condition that

$$\left.\begin{aligned} &\lim_{\eta\to y} \mathscr{J} \frac{\partial}{\partial \eta} \int_0^x \frac{A(t)}{\sqrt{\{[a(t)]^2-\eta^2\}}}\,dt \\ &\lim_{\eta\to y} \mathscr{J} \left[\frac{\partial}{\partial \eta} \int_0^{c_0} \frac{A(t)}{\sqrt{\{[a(t)]^2-\eta^2\}}}\,dt + \frac{\partial}{\partial \eta} \int_{c_0}^x A(t)\, g(t,\eta)\,dt\right] \end{aligned}\right\} = V\alpha \begin{cases} \text{for } 0<x<c_0 \\ \\ \text{for } c_0<x<c \end{cases} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

for any point $(x, y, 0)$ on the aerofoil. It may be noticed that

$$\lim_{\eta\to y} \mathscr{J} \frac{\partial}{\partial \eta} \frac{1}{\sqrt{\{[a(x)]^2-\eta^2\}}} = \lim_{\eta\to y} \mathscr{J} \frac{\partial}{\partial \eta}\, g(x,\eta) = 0$$

at all points $(x, y, 0)$ of the aerofoil, so that by (8) the normal induced velocity is constant along the chords of the aerofoil for *any* choice of $A(t)$. Its magnitude is determined entirely by the values of the integrands in (8) ahead of, and in the region of, the leading edge. It is easy to show that the first equation in (9) is satisfied by

$$A(t) = -V\alpha\, a(t)\, a'(t), \qquad 0<t<c_0 \quad . \quad . \quad . \quad . \quad . \quad (10)$$

In fact,

$$\begin{aligned} \int_0^x [-V\alpha\, a(t)\, a'(t)] \frac{dt}{\sqrt{\{[a(t)]^2-\eta^2\}}} &= -V\alpha \left[\sqrt{\{[a(t)]^2-\eta^2\}}\right]_{t=0}^{t=x} \\ &= V\alpha \left[i\eta - \sqrt{\{[a(x)]^2-\eta^2\}}\right] \end{aligned}$$

and so, for points $(x, y, 0)$ on the aerofoil

$$\lim_{\eta\to y} \mathscr{J} \frac{\partial}{\partial \eta} \int_0^x [-V\alpha\, a(t)\, a'(t)] \frac{dt}{\sqrt{\{[a(t)]^2-\eta^2\}}} = V\alpha\, \mathscr{J} \left[i - \frac{y}{\sqrt{\{[a(x)]^2-\eta^2\}}}\right] = V\alpha$$

as required. It will be seen that the solution implied by (10) agrees with Ref. 1 for the particular case considered here.

The left hand side of the second equation in (9) now becomes, with the specified value of $A(t)$ for $0<t<c_0$,

$$\begin{aligned} &\lim_{\eta\to y} \left[\frac{\partial}{\partial \eta} \int_0^x [-V\alpha\, a(t)\, a'(t)] \frac{dt}{\sqrt{\{[a(t)]^2-\eta^2\}}} + \frac{\partial}{\partial \eta} \int_{c_0}^x A(t)\, g(t,\eta)\,dt\right] \\ &\quad = \lim_{\eta\to y} V\alpha\, \mathscr{J} \left[i - \frac{\eta}{\sqrt{\{[a(c_0)]^2-\eta^2\}}}\right] + \lim_{\eta\to y} \mathscr{J} \frac{\partial}{\partial \eta} \int_{c_0}^x A(t)\, g(t,\eta)\,dt \\ &\quad = V\alpha + \frac{V\alpha y}{\sqrt{(y^2-a_0^2)}} + \lim_{\eta\to y} \mathscr{J} \left[\frac{\partial}{\partial \eta} \int_{c_0}^x A(t)\, g(t,\eta)\,dt\right] \end{aligned}$$

where $a_0=a(c_0)$, and where only points of the aerofoil for which $y>a_0$ are considered.

The condition for $A(t)$, $c_0<t<c$, now becomes

$$\lim_{\eta\to y}\mathscr{I}\left[\frac{\partial}{\partial\eta}\int_{c_0}^{x}A(t)\,g(t,\eta)\,dt\right]=-V\alpha\frac{y}{\sqrt{(y^2-a_0^2)}}\,. \qquad (11)$$

for points $(x, y, 0)$ of the aerofoil such that $x>c_0$, $y>a_0$. The left hand side of (11) may be modified slightly by replacing $\partial/\partial\eta$ by $\partial/\partial y$ since $\eta=y+iz$. Integrating with respect to y on both sides of (11) an alternative condition for $A(t)$ results:—

$$\lim_{\eta\to y}\mathscr{I}\int_{c_0}^{x}A(t)\,g(t,\eta)\,dt=-V\alpha\sqrt{\{y^2-[a(c_0)]^2\}}\,. \qquad (12)$$

More generally, a function of x might be added to the right hand side of (12), but in any case, any solution of (12) leads to a solution of (11). However, the consideration of the limiting case $x\longrightarrow c_0$, $y\longrightarrow a(c_0)$ shows that (12) is the only possible form for the integrated equation. If η tends to y (*i.e.* z tends to 0) on the left hand side of (12) before integration, and then $(x, y, 0)$ tends to the leading edge *for given* y, so that x and y are linked by the relation $y=a(x)$, then

$$\mathscr{I}\int_{c_0}^{x}A(t)\,g(t,y)\,dt=\mathscr{I}\int_{c_0}^{x}A(t)\,g(t,a(x))\,dt$$

$$=\mathscr{I}\int_{c_0}^{x}A(t)\left\{\int_{0}^{a(x)}\left(K'-E'-\frac{[a(t)]^2-[b(t)]^2}{[a(t)]^2-s^2}K'\right)\frac{ds}{\sqrt{\{([a(t)]^2-s^2)(s^2-[b(t)]^2)\}}}\right\}dt$$

where the modulus of the elliptic integrals is $k=b(t)/a(t)$ as before.

$g(t,a(x))$ must be evaluated for values of x not less than t, so that $a(x)\geqslant a(t)$. For this purpose the path of integration of

$$g(t,a(x))=\int_{0}^{a(x)}\left\{K'-E'-\frac{[a(t)]^2-[b(t)]^2}{[a(t)]^2-s^2}K'\right\}\left\{\frac{ds}{\sqrt{\{([a(t)]^2-s^2)(s^2-[b(t)]^2)\}}}\right\}$$

is taken from $s=0$ along the positive imaginary axis to $s=\infty$ and then back along the positive real axis to $s=a(x)$. It has been shown already that $g(t,\eta)$ vanishes for $\eta=\infty$ and so the integral along the positive imaginary axis does not contribute anything to $g(t,a(x))$. Introducing the new variable of integration r by $r=a(t)/s$, $s=a(t)/r$, gives

$$g(t,a(x))=-\int_{0}^{a(t)/a(x)}\left\{K'-E'-\frac{(a^2-b^2)r^2}{a^2(r^2-1)}K'\right\}\frac{dr}{a\sqrt{\{(r^2-1)(1-k^2r^2)\}}}$$

where the argument of a and b is t except where stated otherwise (in the upper limit of the integral). The integrand is pure imaginary in the entire range of integration and so

$$\mathcal{J}g(t,a(x))=-\int_0^{a(t)/a(x)}\left\{K'-E'+\frac{(1-k^2)r^2}{1-r^2}K'\right\}\frac{dr}{a\sqrt{\{(1-r^2)(1-k^2r^2)\}}}. \qquad (13)$$

where the sign $(-)$ is determined by the variation of the square root along the path of integration. Substituting $r=\mathrm{sn}(\tau,k)$ in $N(t,x)=-\mathcal{J}g(t,a(x))$, gives

$$N(t,x)=\frac{1}{a(t)}\int_0^{\mathrm{sn}^{-1}\frac{a(t)}{a(x)}}\{K'-E'+k'^2K'[\mathrm{sc}(\tau,k)]^2\}\,dt$$

$$=\frac{1}{a(t)}\Big[(K'-E')\tau+K'\{\mathrm{dn}(\tau,k)\,\mathrm{sc}(\tau,k)-E(\tau)\}\Big]_0^{\mathrm{sn}^{-1}\frac{a(t)}{a(x)}}$$

$$N(t,x)=\frac{1}{a(t)}\Big[(K'-E')F\left\{\frac{a(t)}{a(x)},\ k\right\}-K'E\left\{\frac{a(t)}{a(x)},\ k\right\}\Big]+$$

$$+K'\frac{\sqrt{\{[a(x)]^2-[b(t)]^2\}}}{a(x)\sqrt{\{[a(x)]^2-[a(t)]^2\}}} \qquad . \quad . \quad . \quad . \quad (14)$$

where F and E are now the incomplete elliptic integrals of the first and second kind respectively. The arguments of F and E are taken to be defined by

$$F\{\beta,k\}=\int_0^{\beta}\frac{dt}{\sqrt{\{(1-t^2)(1-k^2t^2)\}}},\qquad E\{\beta,k\}=\int_0^{\beta}\left\{\frac{1-k^2t^2}{1-t^2}\right\}dt.$$

Equation (12) becomes in terms of $N(t,x)$,

$$\int_{c_0}^{x}A(t)N(t,x)\,dt=V\alpha\sqrt{\{[a(x)]^2-a_0^2\}} \qquad . \quad . \quad . \quad (15)$$

3. Calculation of Chordwise Lift Distribution

By (1) and (7), the pressure difference at a point $(x,y,0)$ of the aerofoil is given by

$$\Delta p=\begin{cases}-2\rho V\dfrac{A(x)}{\sqrt{\{[a(x)]^2-y^2\}}} & \text{for } 0<x<c_0\\[2ex] -2\rho VA(x)\displaystyle\int_b^y\left\{K'-E'-\frac{a^2-b^2}{a^2-s^2}K\right\}\frac{ds}{\sqrt{\{(a^2-s^2)(s^2-b^2)\}}}, & \text{for } c_0<x<c,\ y>0.\end{cases}$$

$$. \quad . \quad . \quad . \quad . \quad . \quad . \quad (16)$$

The lower limit of the integral may be taken as $b(x)$, since the contribution of the integral along the segment $<0,\ b(x)>$ to (5) is imaginary, and also because the Joukowski condition is satisfied at the trailing edge. The modulus of the elliptic integrals is $k=a/b$ where $a=a(x)$, $b=b(x)$ here and elsewhere in the integral.

The lift per unit chord is given by

$$l(x)=\int_{-a(x)}^{a(x)} \Delta p\, dy \qquad \text{for } 0<x<c_0.$$

Thus, in that case

$$l(x)=-2\rho V\int_{-a(x)}^{a(x)} \frac{A(x)\,dy}{\sqrt{\{[a(x)]^2-y^2\}}}=-2\rho VA(x)\left[\sin^{-1}\frac{y}{a(x)}\right]_{-a(x)}^{a(x)}$$
$$=-2\pi\rho VA(x).$$

On taking into account (10),

$$l(x)=2\pi\rho V^2\,\alpha\, a(x)\,a'(x), \qquad 0<x<c_0. \qquad . \quad . \quad (17)$$

On the other hand, for $c_0<x<c$,

$$l(x)=\int_{-a(x)}^{-b(x)}\Delta p\,dy+\int_{b(x)}^{a(x)}\Delta p\,dy=2\int_{b(x)}^{a(x)}\Delta p\,dy$$

$$=-4\rho VA(x)\int_b^a dy\int_b^y\left\{K'-E'-\frac{a^2-b^2}{a^2-s^2}K\right\}\frac{ds}{\sqrt{\{(a^2-s^2)(s^2-b^2)\}}}$$

$$=4\rho VA(x)\int_b^a dy\left\{K'\int_b^y\frac{s^2-b^2}{a^2-s^2}\frac{ds}{\sqrt{[(a^2-s^2)(s^2-b^2)]}}+E'\int_b^y\frac{ds}{\sqrt{\{(a^2-s^2)(s^2-b^2)\}}}\right\}.$$

The first integral in the curly brackets can be transformed, so

$$\int_b^y\frac{(s^2-b^2)^{\frac{1}{2}}}{(a^2-s^2)^{\frac{3}{2}}}ds=\int_b^y\frac{(s^2-b^2)^{\frac{1}{2}}}{s}\frac{s}{(a^2-s^2)^{\frac{3}{2}}}ds$$

$$=\left[\frac{(s^2-b^2)^{\frac{1}{2}}}{s(a^2-s^2)^{\frac{1}{2}}}\right]_b^y+\int_b^y\left[\frac{(s^2-b^2)^{\frac{1}{2}}}{s^2(a^2-s^2)^{\frac{1}{2}}}-\frac{1}{(a^2-s^2)^{\frac{1}{2}}(s^2-b^2)^{\frac{1}{2}}}\right]ds$$

$$=\frac{(y^2-b^2)^{\frac{1}{2}}}{y(a^2-y^2)^{\frac{1}{2}}}-b^2\int_b^y\frac{ds}{s^2(a^2-s^2)^{\frac{1}{2}}(s^2-b^2)^{\frac{1}{2}}}.$$

Then

$$l(x)=4\rho V A(x)\left\{K'\int_b^a \frac{(y^2-b^2)^{\frac{1}{2}}\,dy}{y(a^2-y^2)^{\frac{1}{2}}} - K'b^2\int_b^a dy\int_b^y \frac{ds}{s^2(a^2-s^2)^{\frac{1}{2}}(s^2-b^2)^{\frac{1}{2}}} + \right.$$

$$\left. + E'\int_b^a dy\int_b^y \frac{ds}{(a^2-s^2)^{\frac{1}{2}}(s^2-b^2)^{\frac{1}{2}}}\right\}$$

$$=4\rho V A(x)\left\{K\int_b^a \frac{(y^2-b^2)^{\frac{1}{2}}\,dy}{y(a^2-y^2)^{\frac{1}{2}}} - K'b^2\left[a\int_b^a \frac{ds}{s^2(a^2-s^2)^{\frac{1}{2}}(s^2-b^2)^{\frac{1}{2}}} - \right.\right.$$

$$\left.\left. - \int_b^a \frac{dy}{y(a^2-y^2)^{\frac{1}{2}}(y^2-b^2)^{\frac{1}{2}}}\right] + E'\left[a\int_b^a \frac{ds}{(a^2-s^2)^{\frac{1}{2}}(s^2-b^2)^{\frac{1}{2}}} - \int_b^a \frac{y\,dy}{(a^2-y^2)^{\frac{1}{2}}(y^2-b^2)^{\frac{1}{2}}}\right]\right\}.$$

Putting $y^2=t$, $2y\,dy=dt$ in the first, third and fifth integral, and $s=a\,\mathrm{dn}(\tau,k')$ in the second and fourth, where $k'=\sqrt{(1-k^2)}=\sqrt{\{1-(b/a)^2\}}$, gives

$$l(x)=4\rho V A(x)\left\{\tfrac{1}{2}K'\int_{b^2}^{a^2} \frac{(t-b^2)^{\frac{1}{2}}}{t(a^2-t)^{\frac{1}{2}}}\,dt + \tfrac{1}{2}K'b^2\int_{b^2}^{a^2} \frac{dt}{t(a^2-t)^{\frac{1}{2}}(t-b^2)^{\frac{1}{2}}} - \right.$$

$$\left. - \tfrac{1}{2}E'\int_{b^2}^{a^2} \frac{dt}{(a^2-t)^{\frac{1}{2}}(t-b^2)^{\frac{1}{2}}} - K'k^2\int_0^{K'} \mathrm{nd}^2\,d\tau + E'\int_0^{K'} d\tau\right\}$$

$$=4\rho V A(x)\left\{\tfrac{1}{2}\pi K'(1-k)+\tfrac{1}{2}\pi K'k-\tfrac{1}{2}\pi E'-K'E'+E'K'\right\}$$

or
$$l(x)=2\pi(K'-E')\rho V A(x) \quad . \quad . \quad . \quad . \quad . \quad (18)$$

where the argument of the elliptic integrals is $k=b(x)/a(x)$.

4. Solution of Integral Equation for Small Sweepback of the Trailing Edge

By (18),
$$A(x)=\frac{l(x)}{2\pi\rho V(K'-E')} \quad . \quad . \quad . \quad . \quad . \quad (19)$$

Substituting this expression for A in the integral equation (15) and multiplying by $2\pi\rho V$, gives

$$\int_{c_0}^{x} l(t)\frac{N(t,x)}{K'-E'}\,dt = 2\pi\rho V^2\,\alpha\sqrt{\{[a(x)]^2-a_0{}^2\}} \quad . \quad . \quad . \quad (20)$$

The kernel of (20), $M(t,x)=N(t,x)/(K'-E')$, is always positive as can be seen by the inspection of (13) and (14). If $N(t,x)$ is written in the form

$$N(t,x)=\frac{1}{a(t)}\left[K'\left(F\left\{\frac{a(t)}{a(x)},\ k\right\}-E\left\{\frac{a(t)}{a(x)},\ k\right\}\right)-E'F\left\{\frac{a(t)}{a(x)},\ k\right\}\right]+$$
$$+K'\frac{[a(x)]^2-[b(t)]^2}{a(x)\sqrt{\{[a(x)]^2-[a(t)]^2\}}} \quad . \quad . \quad . \quad . \quad (21)$$

and recalling that $k=b(t)/a(t)$, then as t tends to c_0, for given x, the dominant term in $N(t,x)$ becomes

$$K'\frac{\sqrt{\{[a(x)]^2-[b(t)]^2\}}}{a(x)\sqrt{\{[a(x)]^2-[a(t)]^2\}}}$$

since K' tends to infinity logarithmically as its modulus tends to 0, while the difference $F-E$ tends to 0 as the square of the modulus. Hence

$$\lim_{t\to c_0} M(t,x)=\lim_{t\to c_0}\frac{N(t,x)}{K'-E'}=\frac{1}{a(x)\sqrt{\{[a(x)]^2-a_0^2\}}},\quad \text{for } c_0\leqslant x<c. \qquad (22)$$

On the other hand, $M(t,x)$ becomes infinite of order $1/2$ for given x, as t tends to x.

For any x, x', x'', $c_0\leqslant x'<x''\leqslant x\leqslant c$, by a mean value theorem of the integral calculus,

$$\int_{x'}^{x''} l(t)\,M(t,x)\,dt=\int_{x'}^{x''} l(t)\,M(t,x)\frac{\sqrt{\{[a(x)]^2-[a(t)]^2\}}}{a(t)\,a'(t)}\frac{a(t)\,a'(t)dt}{[a(x)]^2-[a(t)]^2}$$
$$=-l(\xi)\,M(\xi,x)\frac{\sqrt{\{[a(x)]^2-[a(\xi)]^2\}}}{a(\xi)\,a'(\xi)}\left(\sqrt{\{[a(x)]^2-[a(x'')]^2\}}-\right.$$
$$\left.-\sqrt{\{[a(x)]^2-[a(x')]^2\}}\right) \quad . \quad . \quad . \quad (23)$$

where ξ is some intermediate value between x' and x''.

Using (23), the exact value of $l(c_0+0)$ may be found, *i.e.* the limiting value of $l(t)$ as t tends to c_0 from above. Putting $x'=c_0$ and $x''=x$ in (23) and substituting the result on the left hand side of (20), gives

$$l(\xi)\,M(\xi,x)\frac{\sqrt{\{[a(x)]^2-[a(\xi)]^2\}}}{a(\xi)\,a'(\xi)}\sqrt{\{[a(x)]^2-a_0^2\}}=V\alpha\sqrt{\{[a(x)]^2-a_0^2\}}$$

or
$$l(\xi)=\frac{V\alpha\,a(\xi)\,a'(\xi)}{M(\xi,x)\sqrt{\{[a(x)]^2-[a(\xi)]^2\}}}$$

Now let x tend to a_0 from above. Then ξ tends to c_0 also. Furthermore, by (21)

$$M(\xi,x)\sqrt{\{[a(x)]^2-[a(\xi)]^2\}}=\frac{1}{K'-E'}\frac{\sqrt{\{[a(x)]^2-[a(\xi)]^2\}}}{a(\xi)}[K'(F-E)-E'F]+$$
$$+\frac{K'\sqrt{\{[a(x)]^2-[b(\xi)]^2\}}}{a(x)}.$$

and so $$\lim_{\substack{x\to c_0\\ \xi\to c_0}} M(\xi,x)\sqrt{\{[a(x)]^2-[a(\xi)]^2\}}=1$$

by the argument leading up to (22). Hence finally

$$l(c_0+0)=2\pi\rho V^2\alpha a_0 a'(c_0) \quad . \quad . \quad . \quad . \quad (24)$$

so that the lift distribution is continuous across $x=c_0$.

For a simple swallow tail wing (Fig. 1(*b*)), $a'(x)=\text{constant}=a_0/c_0=\tan\gamma$, where γ is the semi apex angle of the wing (Fig. 2). For that case, let $\mu=(c-c_0)/c_0$, then $\tan\gamma$ and μ together determine the aerofoil geometry. The maximum chordwise dimension of the aerofoil then equals $c=(1+\mu)c_0$, the span $d=2(1+\mu)a_0$, the area $S=a_0 c_0(1+\mu)$, and the aspect ratio $=4(1+\mu)\tan\gamma$.

Let $x=(1+\mu s)c_0$, so that x varies between c_0 and c as s varies between 0 and 1. Then $a(x)=a_0(1+\mu s)$ and $b(x)=a_0(1+\mu)s$.

Introducing the variable σ by $t=(1+\mu\sigma)c_0$ gives

$$M^*(\sigma,s)=a_0 M(t,x)=\frac{1}{K'-E'}\left\{\frac{1}{1+\mu\sigma}\left[K'\left(F\left\{\frac{1+\mu\sigma}{1+\mu s},\ k\right\}-E\left\{\frac{1+\mu\sigma}{1+\mu s},\ k\right\}\right)-\right.\right.$$
$$\left.\left.-E'F\left\{\frac{1+\mu\sigma}{1+\mu s},\ k\right\}\right]+\frac{K'}{1+\mu s}\frac{\sqrt{\{(1+\mu s)^2-(1+\mu)^2\sigma^2\}}}{\sqrt{\{(1+\mu s)^2-(1+\mu\sigma)^2\}}}\right\}$$

where the modulus of E and F is $k=k(t)=(1+\mu)\sigma/(1+\mu\sigma)$ and this is also the modulus of E' and K'. Writing

$$l(t)=2\pi\rho V^2\alpha\frac{a_0^2}{c_0}l^*(\sigma)=2\pi\rho V^2\alpha a_0\tan\gamma\, l^*(\sigma)$$

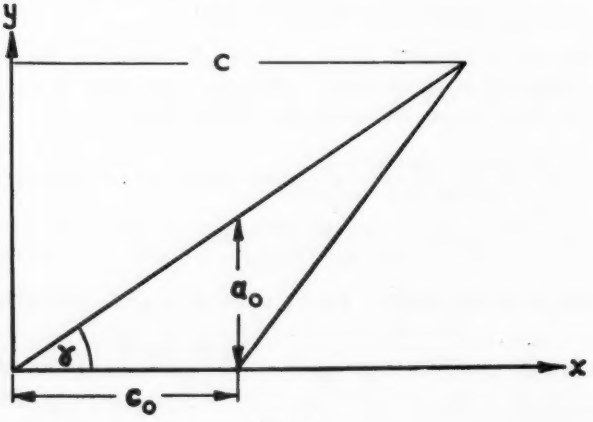

Fig. 2.

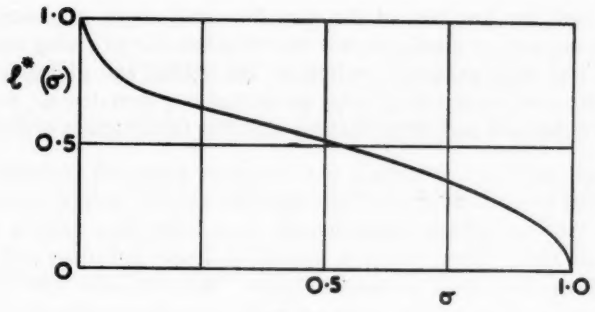

Fig. 3.

equation (20) gives

$$\int_0^s l^*(\sigma)\sqrt{\mu}\,M^*(\sigma,s)\,ds=\sqrt{\left\{\frac{(1+\mu s)^2-1}{\mu}\right\}} \quad . \quad . \quad . \quad (25)$$

For small μ, $k=\sigma+O(\mu)$, and

$$\sqrt{\mu}\,M^*(\sigma,s)=\sqrt{\mu}\left\{K'(\sigma)[K(\sigma)-E(\sigma)]-E'(\sigma)K(\sigma)+K'(\sigma)\frac{\sqrt{(1-\sigma^2)}}{\sqrt{\{2(s-\sigma)\mu\}}}\right\}+O(\mu)$$

$$=\sqrt{\mu}\left\{K'(\sigma)\frac{\sqrt{(1-\sigma^2)}}{\sqrt{\{2(s-\sigma)\mu\}}}-\frac{\pi}{2}\right\}+O(\mu)$$

while

$$\sqrt{\left\{\frac{(1+\mu s)^2-1}{\mu}\right\}}=\sqrt{(2s)}+O(\mu).$$

Equation (25) therefore becomes, in the limit,

$$\int_0^s l^*(\sigma)\frac{K'(\sigma)\sqrt{(1-\sigma^2)}}{K'(\sigma)-E'(\sigma)}\,\frac{ds}{\sqrt{(s-\sigma)}}=2\sqrt{s}. \quad . \quad . \quad . \quad (26)$$

It may be verified directly that this equation is satisfied by

$$l^*(\sigma)=\frac{K'(\sigma)-E'(\sigma)}{K'(\sigma)\sqrt{(1-\sigma^2)}}.$$

$l^*(\sigma)$ is plotted against σ in Fig. 3. As σ tends to 0, $l^*(\sigma)$ tends to 1, in accordance with (24). As σ tends to 1, $K'(\sigma)$ tends to $\pi/2$ while the ratio $[K'(\sigma)-E'(\sigma)]/(1-\sigma^2)$ tends to $\pi/4$. Hence

$$l^*(\sigma)\sim\tfrac{1}{2}\sqrt{(1-\sigma^2)} \qquad \text{as } \sigma\longrightarrow 1.$$

$l^*(\sigma)$ yields the chordwise lift distribution and, using (16) and (18), the pressure

distribution over the rear part of the wing for small angles of sweepback. The result applies not only to simple swallow tail wings but also to a wing such as shown in Fig. 1(*a*), and more generally, so long as the leading and trailing edges of the wing aft of the root chord trailing edge are straight. γ then denotes the (constant) angle between the rear part of the leading edge and the direction of flow.

For larger angles of sweepback (*i.e.* for larger values of μ) the interval $(0, 1)$ may be divided into a number m of sub-intervals, and the integral on the left hand side of (25) may be replaced approximately by a finite sum, using a mean value formula as in (23). There results a system of linear equations with triangular matrix which can be solved without difficulty. However, although a comparison of the results for $m=4$ and $m=5$ seemed to indicate good convergence in a particular case, the general validity of the method is still in doubt.

References

1. Jones, R. T. (1946). Properties of Low Aspect Ratio Pointed Wings at Speeds Below and Above the Speed of Sound. N.A.C.A. Report 835, 1946.
2. Munk, M. M. (1924). The Aerodynamic Forces on Airship Hulls. N.A.C.A. Report 184, 1924.
3. Ribner, H. S. (1947). The Stability Derivatives of Low Aspect Ratio Triangular Wings at Subsonic and Supersonic Speeds. N.A.C.A. Technical Note 1423, 1947.
4. De Young, J. (1950). Spanwise Loading for Wings and Control Surfaces of Low Aspect Ratio. N.A.C.A. Technical Note 2011, 1950.
5. Hodges, M. D. (1950). Nose Controls on Delta Wings. College of Aeronautics Thesis, 1950 (unpublished).
6. Ward, G. N. (1949). Supersonic Flow past Slender Pointed Bodies. *Quarterly Journal of Mechanics and Applied Mathematics*, Vol. II, 1949, pp. 75-97.
7. Heaslet, M. A., Lomax, H. and Spreiter, J. R. (1949). Linearised Compressible Flow Theory for Sonic Flight Speeds. N.A.C.A. Technical Note 1824, 1949.

Correlation Between some Stability Problems for Orthotropic and Isotropic Plates under Bi-Axial and Uni-Axial Direct Stress

W. H. WITTRICK, M.A., Ph.D., A.F.R.Ae.S.

(*Department of Aeronautical Engineering, University of Sydney*)

Summary: Four buckling problems are considered, namely:—

(*a*) A rectangular orthotropic plate, with all edges simply supported, subjected to compression on its ends and a known compression or tension on its sides;

(*b*) the same as (*a*) but with the ends clamped and sides simply supported;

(*c*) a rectangular orthotropic plate, with the ends simply supported and sides clamped, subjected to compression on its ends; and

(*d*) the same as (*c*) but with all edges clamped.

In each case it is shown that the variables involved can all be combined into such a non-dimensional form that a single curve serves to give the value of the end compression required to cause buckling. This curve is identical with the curve of buckling stress coefficient against side ratio for a corresponding isotropic plate under uni-axial compression.

1. Introduction

If a rectangular isotropic plate is subjected to a uni-axial compressive force N_x per unit length, it will buckle when N_x reaches a critical value given by

$$N_x = k\frac{\pi^2 D}{b^2},$$

where k, the buckling stress coefficient, is a function of the side ratio λ and the conditions of edge support. The curves relating k and λ are as shown in Fig. 2 for the four cases where

(*a*) all edges are simply supported,

(*b*) the ends are clamped and the sides simply supported,

(*c*) the ends are simply supported and the sides clamped, and

(*d*) all edges are clamped.

Paper received November 1951.

[The Aeronautical Quarterly, Volume IV, August 1952]

It is the purpose of this paper to show that these curves have a much wider application than to this special case of an isotropic plate under uni-axial compression. It is shown, for example, that the curves corresponding to (*a*) and (*b*), by giving suitable definitions to the non-dimensional parameters k and λ, enable the critical value of N_x for an *orthotropic* plate to be determined when, at the same time, it is subjected to a lateral compressive or tensile force N_y per unit length. Orthotropic plates under uni-axial stress, and isotropic plates under both bi-axial and uni-axial stress appear merely as particular cases of this more general one. Similarly, it is shown that the curves corresponding to (*c*) and (*d*), again by giving suitable definitions to k and λ, enable the critical value of N_x for an orthotropic plate under uni-axial compression to be determined.

Notation

x, y	Cartesian co-ordinates, in directions parallel to the axes of elastic symmetry of the orthotropic plate	*See* Fig. 1
a, b	lengths of the sides of the plate in the x and y directions	
ξ, η	x/a and y/b respectively	
w	lateral deflection of the buckled plate	
N_x, N_y	compressive forces per unit length applied in the x and y directions	
M_x, M_y	bending moments per unit length in the x and y directions	
M_{xy}	twisting moment per unit length	
D_1, D_2, H, K	elastic constants of the orthotropic plate	
D_{12}	$=2H+K$	
D	flexural rigidity of an isotropic plate $=Et^3/[12(1-\sigma^2)]$	
E	Young's modulus of an isotropic plate	
σ	Poisson's ratio of an isotropic plate	
t	thickness of an isotropic plate	
α	$=D_{12}/\sqrt{(D_1D_2)}$	
β	$(a/b)(D_2/D_1)^{\frac{1}{4}}$	
m, n	integers representing the number of half waves of the buckled plate in the x and y directions	
k	a parameter whose value defines N_x	*See* equation (21)
λ	a parameter depending on the ratio of the sides of the plate	
C_1	constant in equation (11)	
C	constant in equation (30)	
ϕ	function defined by equation (27)	

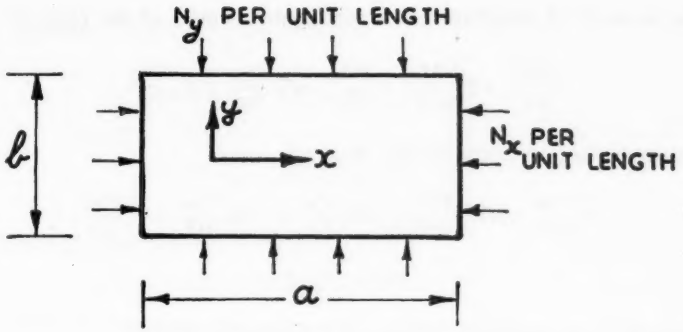

Fig. 1.

2. Basic Equations for an Orthotropic Plate

Consider a flat orthotropic plate subjected to forces N_x and N_y per unit length, as shown in Fig. 1, where the x and y directions are taken to be parallel to the axes of elastic symmetry of the plate.

It can be shown[1] that, if the lateral deflection of the buckled plate at the point (x, y) is w, the moments per unit length are given by

$$\left.\begin{aligned} M_x &= -D_1\frac{\partial^2 w}{\partial x^2} - K\frac{\partial^2 w}{\partial y^2} \\ M_y &= -D_2\frac{\partial^2 w}{\partial y^2} - K\frac{\partial^2 w}{\partial x^2} \\ M_{xy} &= -2H\frac{\partial^2 w}{\partial x \partial y} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (1)$$

where M_x and M_y are the bending moments (producing curvatures in the x and y directions respectively) and M_{xy} the twisting moment.

The constants D_1, D_2, H and K are the elastic constants for the orthotropic plate, and can be obtained by simple tests[2]. In the special case of an isotropic plate they are related by the equations

$$\left.\begin{aligned} D_1 &= D_2 = D \\ K &= \sigma D \\ H &= \tfrac{1}{2}(1-\sigma)D \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where $D = Et^3/[12(1-\sigma^2)]$, is the flexural rigidity of the plate and σ the Poisson's ratio.

The equation of equilibrium of a rectangular element of the plate is[2]

$$\frac{\partial^2 M_x}{\partial x^2}+2\frac{\partial^2 M_{xy}}{\partial x\partial y}+\frac{\partial^2 M_y}{\partial y^2}=N_x\frac{\partial^2 w}{\partial x^2}+N_y\frac{\partial^2 w}{\partial y^2} \quad . \quad . \quad (3)$$

On substituting equations (1) this becomes

$$D_1\frac{\partial^4 w}{\partial x^4}+2D_{12}\frac{\partial^4 w}{\partial x^2\partial y^2}+D_2\frac{\partial^4 w}{\partial y^4}=-N_x\frac{\partial^2 w}{\partial x^2}-N_y\frac{\partial^2 w}{\partial y^2}, \quad . \quad (4)$$

where

$$D_{12}=2H+K. \quad . \quad . \quad . \quad . \quad . \quad (5)$$

It will be seen from equations (2) that for an isotropic plate

$$D_1=D_2=D_{12}=D, \quad . \quad . \quad . \quad . \quad (6)$$

and equation (4) then reduces to its well-known form.

In what follows only rectangular plates are considered whose sides are parallel to the axes of elastic symmetry, and of lengths a and b respectively (Fig. 1). It is convenient to convert x and y to non-dimensional quantities ξ and η by the transformations

$$x=a\xi, \quad y=b\eta \quad . \quad . \quad . \quad . \quad . \quad (7)$$

and to write

$$\left.\begin{aligned}\beta&=\frac{a}{b}\left(\frac{D_2}{D_1}\right)^{\frac{1}{4}}\\ \alpha&=\frac{D_{12}}{\sqrt{(D_1D_2)}}\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Equation (4) then becomes

$$\frac{1}{\beta^4}\frac{\partial^4 w}{\partial\xi^4}+\frac{2\alpha}{\beta^2}\frac{\partial^4 w}{\partial\xi^2\partial\eta^2}+\frac{\partial^4 w}{\partial\eta^4}=-\frac{b^2}{D_2}\left[\frac{N_x}{\beta^2}\left(\frac{D_2}{D_1}\right)^{\frac{1}{4}}\frac{\partial^2 w}{\partial\xi^2}+N_y\frac{\partial^2 w}{\partial\eta^2}\right] \quad . \quad (9)$$

3. Plate with all Edges Simply Supported

In this case the origin of the x and y co-ordinates is taken to be at one corner of the plate so that the equations of the sides are $\xi=0,1$ and $\eta=0,1$.

The boundary conditions are

$$w=M_x=0 \text{ at } \xi=0,1$$

$$w=M_y=0 \text{ at } \eta=0,1$$

On using equations (1) and (7) these reduce to

$$\left.\begin{aligned} w=\frac{\partial^2 w}{\partial \xi^2}=0 \text{ at } \xi=0,1 \\ w=\frac{\partial^2 w}{\partial \eta^2}=0 \text{ at } \eta=0,1 \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (10)$$

The conditions (10) are satisfied by taking

$$w=C_1 \sin m\pi\xi \sin n\pi\eta \quad . \quad . \quad . \quad . \quad . \quad (11)$$

where m and n are integers representing the number of half waves of the buckled plate in the directions of the sides a and b.

Substituting this expression into the differential equation (9) gives

$$\frac{m^4}{\beta^4}+\frac{2\alpha m^2 n^2}{\beta^2}+n^4=\frac{b^2}{\pi^2 D_2}\left[\frac{N_x}{\beta^2}\left(\frac{D_2}{D_1}\right)^{\frac{1}{2}} m^2+N_y n^2\right] \quad . \quad . \quad (12)$$

It will be assumed that the value of N_y is known, so that equation (12) gives N_x in terms of the as yet unknown integers m and n. Thus

$$\frac{b^2 N_x}{\pi^2\sqrt{(D_1 D_2)}}=\frac{m^2}{\beta^2}+2\alpha n^2+\frac{\beta^2 n^4}{m^2}-\frac{b^2 N_y}{\pi^2 D_2}\frac{\beta^2 n^2}{m^2} \quad . \quad . \quad . \quad (13)$$

It now remains to find values for m and n such that N_x has its smallest possible value.

Differentiating equation (13) with respect to n gives

$$\frac{b^2}{\pi^2\sqrt{(D_1 D_2)}}\frac{\partial N_x}{\partial n}=\frac{2\beta^2 n}{m^2}\left[2\left(n^2+\frac{\alpha m^2}{\beta^2}\right)-\frac{b^2 N_y}{\pi^2 D_2}\right] \quad . \quad . \quad (14)$$

Since m and n are integers they must both be greater than or equal to one. It follows therefore that if

$$N_y<\frac{2\pi^2 D_2}{b^2}\left(1+\frac{\alpha}{\beta^2}\right)$$

i.e. if

$$N_y<2\pi^2\left(\frac{D_2}{b^2}+\frac{D_{12}}{a^2}\right) \quad . \quad . \quad . \quad . \quad . \quad (15)$$

the value of $\partial N_x/\partial n$ is positive, so that N_x increases with n, for all possible values of β and m. For N_x to be as small as possible we must then have

$$n=1 \quad . \quad . \quad . \quad . \quad . \quad (16)$$

From now on it will be assumed that if a lateral stress N_y is applied, it

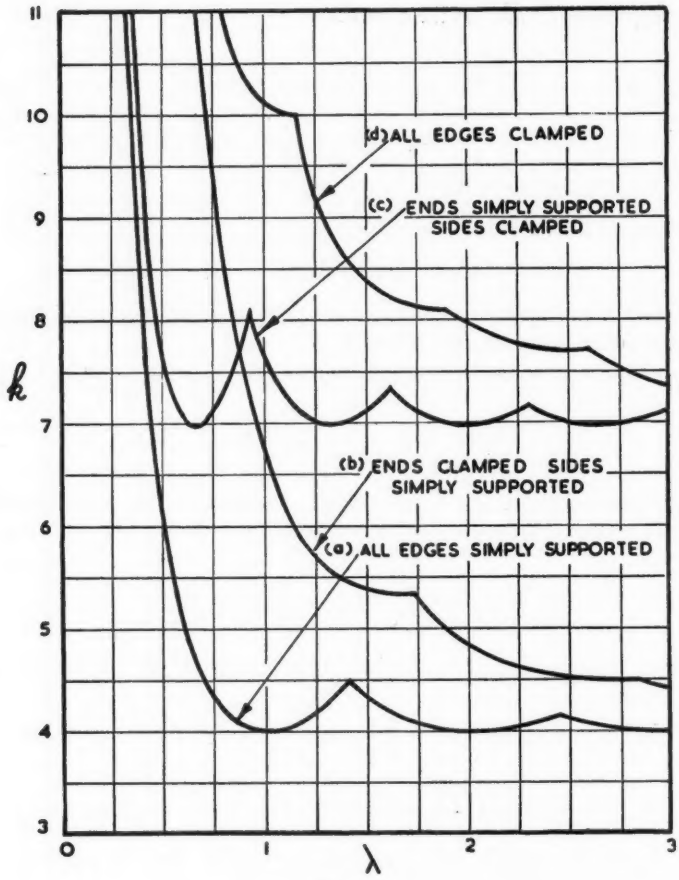

Fig. 2.

satisfies the inequality (15). Equation (16) then holds, and from equation (13)

$$\frac{b^2N_x}{\pi^2\sqrt{(D_1D_2)}} = 2\alpha + \frac{m^2}{\beta^2} + \frac{\beta^2}{m^2}\left(1 - \frac{b^2N_y}{\pi^2D_2}\right) \quad . \quad . \quad . \quad (17)$$

It is seen immediately that if $N_y > \pi^2\dfrac{D_2}{b^2}$ . . . . . (18)

the term in brackets in equation (17) is negative, in which case N_x increases with m. It then follows that $m=1$ and

$$N_x = \frac{\pi^2\sqrt{(D_1D_2)}}{b^2}\left[2\alpha + \frac{1}{\beta^2} + \beta^2\left(1 - \frac{b^2N_y}{\pi^2D_2}\right)\right].$$

On substituting for α and β from equations (8) this gives

$$N_x = \frac{\pi^2}{b^2}\left(2D_{12} + \frac{b^2}{a^2}D_1 + \frac{a^2}{b^2}D_2\right) - \frac{a^2}{b^2}N_y \quad . \quad . \quad . \quad (19)$$

For values of N_y satisfying the inequality

$$N_y < \pi^2 \frac{D_2}{b^2} \quad . \quad . \quad . \quad . \quad . \quad (20)$$

(including all negative, *i.e.* tensile, values) the substitutions

$$\left.\begin{aligned} k &= 2 + \frac{\frac{b^2}{\pi^2}N_x - 2D_{12}}{\left[D_1 D_2\left(1 - \frac{b^2 N_y}{\pi^2 D_2}\right)\right]^{\frac{1}{2}}} \\ \text{and} \qquad \lambda &= \frac{a}{b}\left[\frac{D_2}{D_1}\left(1 - \frac{b^2 N_y}{\pi^2 D_2}\right)\right]^{\frac{1}{4}} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (21)$$

are made.

Equation (17) then gives

$$k = \left(\frac{m}{\lambda} + \frac{\lambda}{m}\right)^2 \quad . \quad . \quad . \quad . \quad . \quad (22)$$

If k is plotted against λ for various integral values of m and the lower envelope is taken, the curve marked (*a*) in Fig. 2 is obtained.

Considering now the definitions of k and λ given in equations (21), it is seen first that for an isotropic plate, represented by equations (6), under compression in the x direction only ($N_y = 0$),

$$k = \frac{b^2 N_x}{\pi^2 D}, \qquad \lambda = \frac{a}{b} \quad . \quad . \quad . \quad . \quad . \quad (23)$$

Thus the curve (*a*) given in Fig. 2 is identical with that relating the buckling stress coefficient and side ratio for a simply supported isotropic plate under compression in one direction.

For an isotropic plate, buckling under stress N_x and a known stress N_y, where according to (20),

$$N_y < \frac{\pi^2 D}{b^2}$$

equations (21) give

$$\left.\begin{aligned} k &= 2 + \frac{\frac{b^2 N_x}{\pi^2 D} - 2}{\left(1 - \frac{b^2 N_y}{\pi^2 D}\right)^{\frac{1}{2}}} \\ \lambda &= \frac{a}{b}\left(1 - \frac{b^2 N_y}{\pi^2 D}\right)^{\frac{1}{4}} \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (24)$$

For an orthotropic plate under compression in the x direction only ($N_y=0$),

$$\left.\begin{aligned} k&=\frac{b^2N_x}{\pi^2\sqrt{(D_1D_2)}}+2\left(1-\frac{D_{12}}{\sqrt{(D_1D_2)}}\right)\\ \lambda&=\frac{a}{b}\left(\frac{D_2}{D_1}\right)^{\frac{1}{4}}\end{aligned}\right\} \quad . \quad . \quad . \quad (25)$$

These three cases, represented by equations (23), (24) and (25), are merely particular cases of the general equations (21). The important point is that in all of them the relation between k and λ is represented by the single curve (a) given in Fig. 2.

Returning to the general equations (21), it is obvious from Fig. 2 that as $\lambda\longrightarrow\infty$, $k\longrightarrow 4$. Thus, for an infinite orthotropic strip, with a side force N_y satisfying the inequality (20),

$$N_x=\frac{2\pi^2}{b^2}\left[D_{12}+\left\{D_1D_2\left(1-\frac{b^2N_y}{\pi^2D_2}\right)\right\}^{\frac{1}{2}}\right] \quad . \quad . \quad . \quad (26)$$

4. Plate with Clamped Ends and Simply Supported Sides

Here the origin of the x and y co-ordinates is taken to be as shown in Fig. 3, so that the equations of the sides are $\xi=\pm\frac{1}{2}$, $\eta=0, 1$.

The boundary conditions are

$$(i)\quad w=\frac{\partial w}{\partial\xi}=0 \text{ at } \xi=\pm\tfrac{1}{2}$$

$$(ii)\quad w=\frac{\partial^2 w}{\partial\eta^2}=0 \text{ at } \eta=0, 1.$$

The conditions (ii) are satisfied by taking a mode of the form

$$w=\phi(\xi)\sin\pi\eta \quad . \quad . \quad . \quad . \quad . \quad (27)$$

It will be seen that it is assumed that the plate buckles with only one half wave in the y direction. This is known to be true when $N_y=0$ and it will certainly be true when $N_y<0$, *i.e.* tensile. As N_y increases positively, it is to be expected that at some value the mode will change from one to two half waves in the y direction. In what follows, however, a limit is placed on the upper value of N_y by the inequality (20), and it is assumed that, if this is satisfied, the mode containing one half wave in the y direction gives the lowest value for N_x.

Substituting equation (27) into the differential equation (9) gives

$$\frac{1}{\beta^4}\frac{d^4\phi}{d\xi^4}+\frac{\pi^2}{\beta^2}\left(\frac{b^2N_x}{\pi^2\sqrt{(D_1D_2)}}\right)-2\alpha\frac{d^2\phi}{d\xi^2}+\pi^4\left(1-\frac{b^2N_y}{\pi^2D_2}\right)\phi=0\,.$$

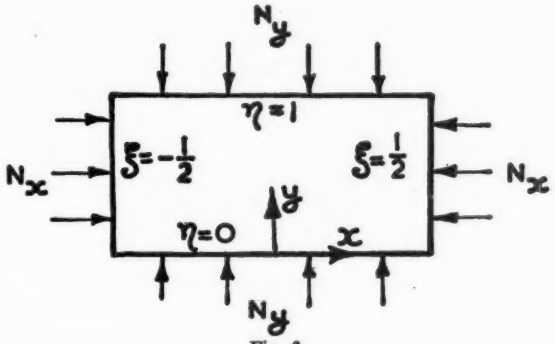

Fig. 3.

Defining k and λ by equations (21), assuming the inequality (20) to be satisfied, this differential equation for ϕ becomes

$$\frac{1}{\lambda^4}\frac{d^4\phi}{d\xi^4}+\frac{\pi^2}{\lambda^2}(k-2)\frac{d^2\phi}{d\xi^2}+\pi^4\phi=0 \quad . \quad . \quad . \quad . \quad (28)$$

The boundary conditions (*i*) when written in terms of ϕ by equation (27), become

$$\phi=\frac{d\phi}{d\xi}=0 \text{ at } \xi=\pm\tfrac{1}{2} \quad . \quad . \quad . \quad . \quad (29)$$

Now it has already been shown that for an isotropic plate under compression in the x direction only the parameters k and λ reduce to equations (23). Thus the differential equation (28) and its boundary conditions (29) are identical with those for the isotropic plate under compression in one direction. It follows that in the orthotropic plate, with both N_x and N_y applied, the relation between k and λ is identical with that for the isotropic plate with only N_x applied. This relation, which has been worked out by F. Schleicher and his results quoted by Timoshenko[2], is shown by curve (*b*) in Fig. 2.

Thus, once again, the cases represented by equations (23), (24) and (25) are merely particular cases of the more general equations (21), and in all of them the relation between k and λ is represented by the single curve (*b*) given in Fig. 2.

5. Other Types of Edge Support

It has been seen that if all edges are simply supported, or if the ends are clamped and the sides simply supported, it is possible to combine the variables into the non-dimensional forms k and λ in such a way that a single curve of k against λ serves to specify the buckling load of an orthotropic plate under bi-axial direct stress, as long as the lateral stress does not exceed a certain value. If other types of edge support are considered, it is found that theoretically this is no longer possible

and that families of curves are required. These curves, each one corresponding to a different value of α, have been derived by Smith[1] for the case of uni-axial compression when the sides are clamped and the ends simply supported, and when all edges are clamped. It appears from his results, however, that if k is defined by the equation

$$k=\frac{b^2N_x}{\pi^2\sqrt{(D_1D_2)}}+C\left(1-\frac{D_{12}}{\sqrt{(D_1D_2)}}\right) \quad . \quad . \quad . \quad (30)$$

where $C=2.4$ for the case of ends clamped and sides simply supported, and $C=2.46$ for the case of all edges clamped, then the curves of k against λ, where λ is given by the second of equations (25), almost exactly coincide with that for the corresponding isotropic plate. The maximum error involved in this assumption never exceeds one per cent., and for most practical applications it would amount to less than $\frac{1}{2}$ per cent. The appropriate isotropic curves (*c*) and (*d*) are given in Fig. 2.

6. Conclusions

It has been shown that a single curve relating the parameters k and λ serves to specify the buckling load N_x in each of the following cases.

(*a*) Orthotropic plate, with all edges simply supported, subjected to compressive forces N_x and N_y on its edges, where N_y is known and is less than π^2D_2/b^2 (including all tensile values). In this case k and λ are given by equations (21).

(*b*) The same as (*a*) but with the ends clamped and the sides simply supported.

(*c*) Orthotropic plate with the ends simply supported and the sides clamped, under uni-axial compression N_x. k is defined approximately by equation (30) with $C=2.4$, and λ by the second of equations (25).

(*d*) Orthotropic plate with all edges clamped, under uni-axial compression N_x. k is defined approximately by equation (30) with $C=2.46$, and λ by the second of equations (25).

In each of these the corresponding isotropic plate appears as a special case simply by using equations (6) in the expressions for k and λ.

REFERENCES.

1. SMITH, R. C. T. (1944). The Buckling of Flat Plywood Plates in Compression. Australian Report ACA-12, 1944.

2. TIMOSHENKO, S. (1936). *Theory of Elastic Stability*. McGraw-Hill, 1936.

The Dust Problem in Hot-Wire Anemometry

D. C. COLLIS

(Fluid Motion Laboratory, University of Manchester; now at Aeronautical Research Laboratory, Melbourne)

Summary: Experimental work is described showing that a major cause of changes in the calibrations of hot-wire anemometers consists in the precipitation of dust on the upstream surface of the wire during use. A simple means of cleaning the air in closed-circuit wind tunnels is described. The rate of change of calibration due to dust was reduced in the Manchester University Fluid Motion Laboratory by at least a factor of 15, by adequate cleaning of the air and by minimising the amount of interchange between the tunnel air and the atmosphere. It is concluded that where accurate results are required from hot wires used in a wind tunnel, the tunnel should be designed to be dust-free.

1. Introduction

A serious difficulty in the use of the hot-wire anemometer is the instability of its calibration. This has been attributed to changes in the physical structure of the wire, variations in the resistance of the junctions between the wire and its supports, and the accumulation of dust on the surface of the wire. The first two of these troubles can be alleviated by annealing the wire before calibration and by improved soldering or welding techniques. The dust problem is more difficult, as hot wires are normally operated in atmospheric air which is never free from dust. There is a progressive accumulation of dust on the wire, with a corresponding change in the heat-loss characteristics. As the hot wire is a delicate structure, it cannot be cleaned satisfactorily without distorting or breaking it. For the measurement of small steady velocities, the problem has been solved satisfactorily by Simmons[1], by enclosing the wire in a silica tube of relatively large diameter. The rate of heat transfer from the large cylinder is less sensitive to the presence of small dust particles than is the rate of transfer from a fine wire, and the silica tube may be safely cleaned at suitable intervals. For the measurement of rapid fluctuations of velocity Simmons' solution is not applicable, because the device has a considerable time lag. This paper presents evidence that, in this application at least, dust deposits on the wire are responsible for the most serious changes in calibration. A simple method is described, applicable to closed-circuit wind tunnels, by which the dust concentration in the air can be reduced to a negligible level.

Paper received March 1951.

[The Aeronautical Quarterly, Volume IV, August 1952]

Notation

a	radius of a dust particle
n	number of dust particles per c.c. of air
I	hot-wire anemometer heating current
K	constant in equation (1)
r	a representative length, radius of a wire or cylinder
R	resistance of anemometer wire when heated by current I
R_a	resistance of anemometer wire at the temperature of the air
U	velocity of the undisturbed air stream
η	dust-collecting efficiency of a cylinder
μ	denotes one micron (10^{-6} metre)
ν	kinematic viscosity of air
σ	density of a dust particle
ρ	density of air

2. The Wind Tunnel

The wind tunnel used, which has not previously been described, is of the closed-circuit type, driven by a 6 h.p. variable-speed A.C. motor (Fig. 1). The working section is 50 cm. square and is vented to the atmosphere. The maximum wind speed is 38 metres/sec. and the contraction ratio is 20:1. Screens are provided for the reduction of turbulence (Fig. 1). The longitudinal and lateral components of the turbulence in the working section are 0.02 and 0.03 per cent.

3. The Effect of Dust

In attempting to compare the relative merits of constant-resistance and constant-current calibration of hot wires, it was immediately found that the inconstancy of results obtained by either method obscured any difference which might exist between the two. Experience at the Aeronautical Research Laboratory, Melbourne, suggested that the changes in calibration were probably caused by dust. At Melbourne, a diminution of wire sensitivity generally occurred at the rate of about $\frac{1}{2}$ per cent. per minute during the first half-hour of running at a mean velocity of

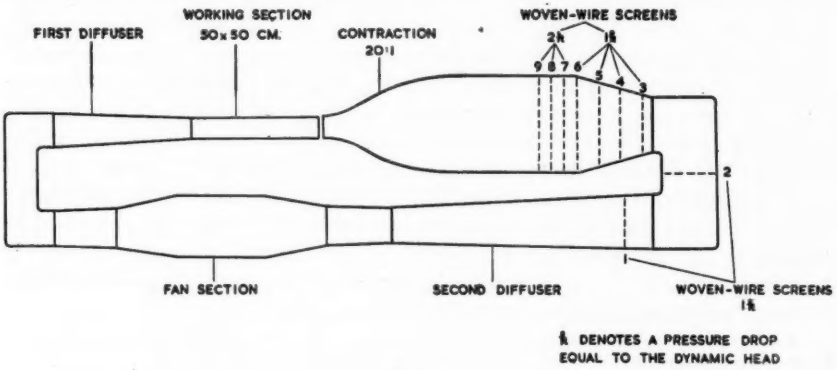

Fig. 1.

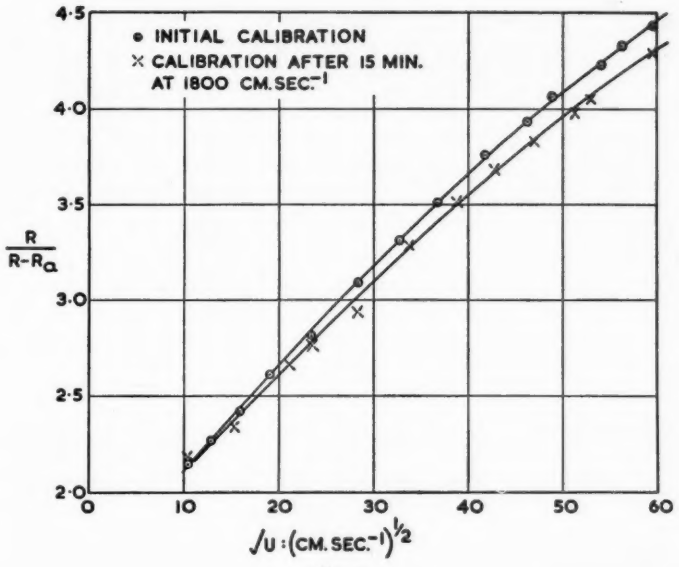

Fig. 2.

30 metres/sec. This was accompanied by a slight increase in the time constant M of the wire (defined by Dryden and Kuethe[2]). The sensitivity of a particular wire which was exposed to a 30 metres/sec. air stream for about $1\frac{1}{2}$ hours changed by more than 25 per cent. The frequency response of this wire changed radically, so that it was no longer possible to compensate the wire according to the law given by Dryden and Kuethe. Microscopic examination of the wire revealed a thick encrustation of dust on the upstream side of the wire. On washing the wire in water and several organic solvents most of the dust was removed. The sensitivity was then found to be restored almost to its original value and the frequency response again assumed the theoretical form, with the time constant M close to its original value.

In the present work at Manchester, microscopic examination of the wires showed dust deposits up to 10 microns (μ) in thickness, but these were not appreciably affected by common solvents. Most of the dust particles were considerably smaller than the wire diameter (2.5 μ) and the smallest particles observed appeared to be about 0.2 μ in diameter. At a later stage, when much of the dust had been extracted from the air in the tunnel, the deposits on the wires were found to consist chiefly of transparent globules, resembling drops of oil. These globules were only dissolved very slowly by benzine but could be evaporated from the middle of the wire by heating it electrically.

The results of some heat-transfer measurements made at Manchester are given in Figs. 2 and 3. They illustrate the magnitude of the changes produced

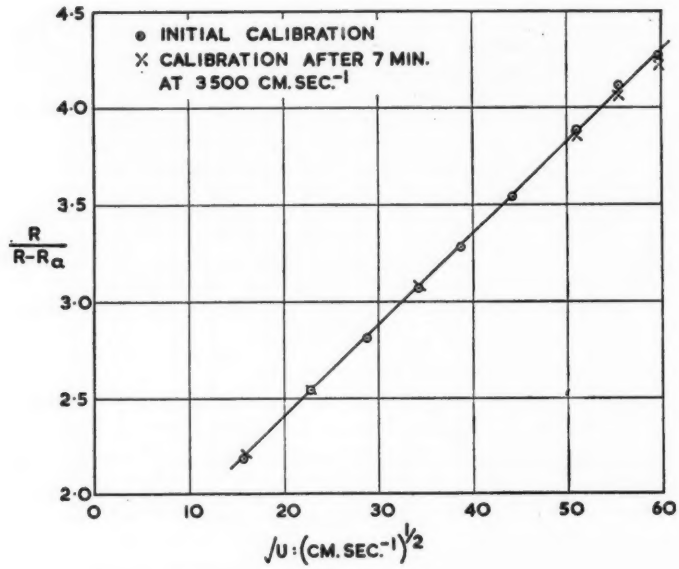

Fig. 3.

by dust and the improvement which can be effected by cleaning the air. The slope of the curve of $R/(R-R_a)$ at a given point is a measure of the hot-wire sensitivity at that particular operating condition. The calibration curves of Fig. 2 were taken in uncleaned air with a running period of 15 min. at 18 metres/sec. between the two calibrations. The rate of change of sensitivity was estimated to be roughly one per cent. per minute in this case. The calibration points shown in Fig. 3 were measured with a new hot wire, after the tunnel had been partially freed from dust. The initial points lie close to a straight line, and those taken after 7 minutes' further running at 35 metres/sec. deviate only slightly from the same line at the high-velocity end. Although the wires used in the measurements of Figs 2 and 3 were subjected to approximately the same mass flow of air, the change in sensitivity was much less in the cleaner air. At a later stage in the cleaning process it was possible to operate a wire for 30 to 40 minutes at 30 metres/sec. with not more than 2 per cent. change in sensitivity. Under the microscope only a layer of transparent globules was observed after the completion of the run. It is believed that if the interchange of the tunnel air with that of the atmosphere could be reduced it would be possible to remove much of this oily matter, with a further improvement in stability of the hot-wire calibrations.

The heat-transfer measurements were not extended to zero velocity, but Betchov[3] reports that the effect of dust is to increase the heat transfer at zero velocity. He remarks that the sensitivity of the wire to velocity changes is unaltered, which is certainly not true of the results given here.

At various stages of the work photo-micrographs were taken of the dust deposits, some of which are reproduced in Fig. 4. The magnification varies considerably because of the necessity of using different microscopes. The deposits consist of particles appreciably smaller than the wire diameter (the estimated particle size is $0.2\,\mu$ to 1.2μ), which are almost exclusively attached to the front surface of the wire. This is in disagreement with Simmons[1], who states that on fine wires the particles form a fairly uniform layer over the whole surface—a configuration only likely to occur if the wire is left exposed to random air currents when not in use.

The hot wires used in this work were platinum of nominal diameter $2.5\,\mu$, prepared from 10:1 Wollaston wire. As in most cases the ambient air temperature rose a few degrees during the course of a run, the value of R_a was adjusted accordingly before computing $R/(R-R_a)$.

(*a*) After 15 minutes operation at 18 metres/sec. Air not cleaned.

(*b*) After 15 minutes operation at 30 metres/sec. Air not cleaned.

(*c*) After 90 minutes operation at 30 metres/sec. Air not cleaned. Wire broken, causing dust layer to peel off.

(*d*) After 90 minutes operation at 30 metres/sec. Air cleaned. Mostly transparent droplets and one unusually large particle.

(*e*) After 90 minutes operation at 30 metres/sec. Air cleaned. Transparent droplets only.

Fig. 4.

Photo-micrographs of hot wires. The upstream surface of each wire is uppermost in the photographs. The diameter of the etched portion of the wires is nominally 2.5μ.

4. Removal of Dust from the Wind Tunnel

From a study of research papers on atmospheric pollution of British cities[4, 5] it is concluded that the majority of the particles deposited on hot wires consist of ash and solid combustible matter, probably mostly carbon, while the transparent globules are probably tar.

Several measures were tried to eliminate the dust from the air in the wind tunnel or at least to reduce the concentration to a negligible level. These are described, with emphasis on the mechanical method which finally proved successful.

4.1. Oiled wire screens

This, the most obvious method, has probably been exploited in most wind tunnels where the hot wire has been used for extensive quantitative measurements. Modern low-turbulence wind tunnels always have a number of woven-wire screens installed in the settling chamber upstream of the working section. Application of a film of oil to such screens results in the collection of a considerable quantity of dust and other foreign matter. In this particular instance, screens Nos. 1 and 2 (Fig. 1), were sprayed with oil and paraffin respectively. Examination after many hours operation showed deposits of dust and organic matter (wood particles, insects, and hairs). That this precipitated matter could consist only of relatively large particles is shown later. It was confirmed by the fact that at no stage in the cleaning process was a reduction observed in the rate of deposition of dust on the hot wires. The process serves a useful purpose, however, since the presence of large particles in the air stream results in frequent breakage or distortion of the fragile hot wires. Even untreated screens collect dust on the upstream side, but the addition of the oil enables a larger deposit to be held on the wire and facilitates subsequent cleaning of the screens with paraffin or benzine. The oil must be limited to a very thin film as any surplus will be blown off the screen and will then precipitate on the hot wires.

4.2. Electrostatic precipitation

The electrostatic type of dust precipitator used in industry acts efficiently on particles as small as 0.2 μ. All the air in the wind tunnel could be circulated periodically through such a device, but this would add considerably to the cost and complexity of what is essentially a simple item of equipment. Nevertheless, it was thought that a simplified application of the electrostatic method might be profitable, particularly as the recirculation of the air allows the use of a low-efficiency device. The experiment consisted of charging screen No. 2 (Fig. 1) to a potential of 3,000 volts and operating the tunnel at 18 metres/sec. for 6 hours. Microscopic examination of the hot wires at this stage indicated a definite reduction in the dust content of the air, but repeated calibrations of a single wire still showed large discrepancies.

To improve this method by raising the potential of the screen, it would have been necessary to insulate the screen from the wooden structure of the tunnel. Two further considerations, the provision of a high-voltage power supply with protective

devices and the necessity for avoiding the production of corrosive nitrogen peroxide, led to the abandonment of the electrostatic technique in favour of the simpler mechanical process now described.

4.3. Mechanical process for extraction of dust

This process is essentially a modification of that described in Section 4.1. To explain why the woven-wire screens in the low-speed sections of the tunnel do not collect small dust particles like the hot wire, it is convenient to refer to an analysis by Taylor[6] in a paper dealing with the collection of water droplets on the surface of a body in an air stream. By considering the equations of motion of an airborne particle in the neighbourhood of an obstacle and assuming that the drag of the particle is given by Stokes' law, the following similarity condition can be derived:—

$$\frac{2}{9}\left(\frac{\sigma}{\rho}\right)\left(\frac{a^2}{L^2}\right)\left(\frac{LU}{\nu}\right)=K. \qquad . \quad . \quad . \quad . \quad (1)$$

The equations of motion may be solved by step-by-step integration and thus the particle paths can be plotted. Taylor makes the calculation for the region near a stagnation point and shows that no particles reach the surface of a cylinder of diameter $2L$ if $K < \frac{1}{8}$. Using the equations given by Taylor, Glauert[7] has computed particle paths in the vicinity of cylinders and aerofoils by step-by-step integration for different values of K.

The work described refers to flow around obstacles at high Reynolds number. It is assumed in the following discussion that the results are applicable even at the low Reynolds number of the flow around fine wires. Experience shows that the deductions made are correct, at least qualitatively, and it was not considered worth repeating the analyses of Taylor and Glauert for the low Reynolds number case.

If all particles of radius a, approaching a cylinder of radius L in a band of width $2L$, are deposited, the rate of deposition per unit length of cylinder is $2nLU$ where n is the number of particles per c.c. of air. The collecting efficiency η of the cylinder, for particles of radius a, may therefore be defined as the ratio of the number of particles deposited per unit length per second to the quantity $2nLU$.

The following table shows values of K and η for various wires and for a typical particle of diameter 0.8μ and density 2 gm./cm.^{3} The velocity is the maximum attainable at the section of the tunnel where the wire is usually located. Also shown is the minimum size of particle ($2a_{min}$) collected.

Description	$2L$ (μ)	U cm. sec.$^{-1}$	K	η per cent.	$2a_{min.}$ (μ)
Wire Gauze Nos. 1, 2	500	500	0.083	0	1.0
Wire Gauze Nos. 7, 8, 9	200	150	0.061	0	1.1
Wire Gauze	200	1800	0.75	30	0.33
Wire Grid	100	3000	2.51	70	0.18
Wollaston Wire Unetched	25	3000	10	90	0.09
Wollaston Wire Etched	2.5	3000	100	100	0.03

The wire gauzes in the low-speed sections of the tunnel are seen to be totally ineffective in collecting particles smaller than 1.0 μ, whereas the hot wire collects particles as small as 0.03 μ. It appears possible, however, to choose conditions so that a wire screen will be effective as a precipitator. The requirements, as indicated by equation (1), are maximum air velocity through the screen and minimum size of wire. In addition, the total projected area of the wires should be large, but not so large that the maximum speed of the tunnel is seriously reduced. The ratio of the projected area of the wires to the area of the gauze is usually termed the blockage of the gauze. The screen may consist of a number of parallel wires, commercial woven-wire gauze, or an open-weave fabric. Several arrangements have been tried in the working section of this tunnel. The materials were mounted on wooden frames which could be inserted quickly into the aperture normally used for turbulence-producing grids.

(*a*) *Woven-wire gauze*

The finest commercial gauze available with sufficiently open weave is that used in screens Nos. 7, 8, 9, *i.e.* 200 μ wires and mesh size 0.85 mm., giving a blockage of 42 per cent. With such a gauze in the working section of the tunnel the maximum speed is reduced from 38 metres/sec. to 18 metres/sec. The theoretical collecting efficiency of a single 200 μ wire at this speed is given in Fig. 5 as a function of particle size. This curve may be used to calculate the time required to reduce the concentration of particles of a given size in the whole tunnel (*i.e.* 67 cubic metres of air) to say one per cent. of their initial concentration, assuming no interchange of air between the tunnel and the outside air. Obviously, this time is large for particles less than 0.4 μ in diameter, but experience showed that most of the troublesome dust was removed in 8 to 10 hours. Two aspects of the performance of the gauze which are not readily estimated are the number of particles which strike the wires and fail to adhere, and the error introduced by assuming the gauze to behave as a group of cylinders.

(*b*) *Fabrics*

Most of the woven fabrics examined proved to have too high an air resistance. A very open muslin of mesh size approximately 1.25 mm. has been used with some success. The thread diameter was very irregular, with an average value of about 150μ. The resistance and collecting efficiency are similar to that of the 200 μ wire gauze. The only other materials which show promise are a bolting silk of mesh size 0.88 mm. and thread diameter about 180 μ, and a lace tulle with mesh size 1.92 mm. formed from silk threads about 60 μ in diameter.

(*c*) *Parallel-wire grid*

The device in use at present consists of a double array of 100 μ wires formed by winding a flat coil over a square metal framework. The effective blockage of the grid is only 12½ per cent., and could be doubled with advantage if constructional difficulties permitted. The maximum air speed which can be maintained through

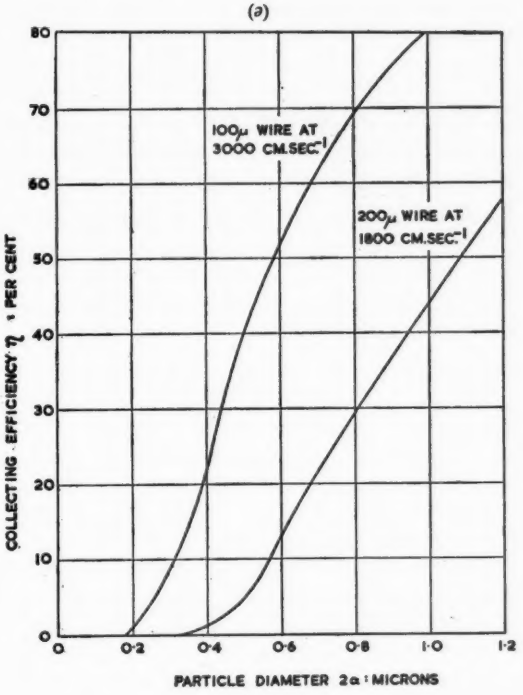

Fig. 5.

the grid is about 30 metres/sec. The theoretical collecting efficiency of a single 100 μ wire at this speed is shown in Fig. 5, and is greatly superior to a single wire (diameter 200 μ) of the gauze described in (*a*). In comparing the overall efficiencies of the grid and the gauze, it is necessary to take into account the total projected areas of wire and the relative rates at which the air in the tunnel is re-circulated at the different speeds. Allowing for these effects, the parallel-wire grid is found to have a greater efficiency for all particle sizes below 0.9 μ. No quantitative observations of dust concentration in the tunnel have been made but the experience gained so far with this grid is in qualitative agreement with the foregoing remarks.

5. Limitations to Removal of Dust from the Wind Tunnel

Two factors set a lower limit to the dust concentration that can be achieved. During operation of the tunnel air escapes through leaks in the return duct, and an equal quantity of contaminated air enters the tunnel at the working section. This effect was reduced in the present work by taking all possible steps to seal the leaks in the return duct. Also, because of the compression of the air in the return duct, there is an intake of contaminated air each time the air stream is accelerated. Thus,

unless elaborate alterations are made to the tunnel there is an ultimate limit to the purity of the air stream.

6. Conclusions

The instability of calibration of hot-wire anemometers is caused mainly by the precipitation of minute dust particles on the upstream surface of the wire while it is exposed to an air stream. In Great Britain this dust is derived from coal-smoke.

The rate of change of hot-wire sensitivity depends on local atmospheric conditions but may easily be as much as one per cent. per minute at moderate air speeds. In a closed wind tunnel used in an industrial area it has been possible to reduce the dust content of the air and diminish the effect of dust on stability of calibration by at least 15 times.

Special attention should be given to the design of wind tunnels where hot wires are to be used, so that the interchange of the air in the wind tunnel with that of the atmosphere can be kept to a minimum. This involves designing a structure which will not develop leaks through repeated stressing of the walls, and careful consideration should be given to the positioning of the vent holes. Some means of cleaning the air must be provided. Since the air is re-circulated this device need only be of low efficiency, but it must be effective on particles much less than $1.0\,\mu$ in diameter. In high-speed wind tunnels the problem is more serious; there is a greater tendency to develop leaks, and some dust may be produced by abrasion of the walls, particularly if these are of concrete.

Acknowledgments

The author wishes to acknowledge the assistance provided by frequent discussions with Mr. W. A. Mair. During the course of the work the author was holder of a studentship provided by the Department of Supply (Australia).

References

1. Simmons, L. F. G. (1949). A Shielded Hot-Wire Anemometer for Low Speeds. *J. Sci. Instrum. Phys. Ind.*, Vol. 24, pp. 407-11. December 1949.
2. Dryden, H. L. and Kuethe, A. (1929). Measurement of Fluctuations of Air Speed by a Hot-Wire Anemometer. N.A.C.A. Report 320, 1929.
3. Betchov, R. (1949). Théorie Non-Linéaire de l'Anémomètre à Fil Chaud. *Koninklijke Nederlandsche Akademie van Wetenschappen.* Vol. LII, No. 3, 1949.
4. — (1945). Atmospheric Pollution in Leicester. Department of Scientific and Industrial Research, Atmospheric Pollution Research, Technical Paper No. 1, 1945.
5. Shaw, N. and Owens, J. S. (1925). *The Smoke Problem of Great Cities.* Constable, London, 1925.
6. Taylor, G. I. (1940). Notes on Possible Equipment and Technique for Experiments on Icing on Aircraft. R. & M. 2024, 1940.
7. Glauert, Muriel (1940). A Method of Constructing the Paths of Raindrops of Different Diameters Moving in the Neighbourhood of (1) a Circular Cylinder, (2) an Aerofoil, Placed in a Uniform Stream of Air; and a Determination of the Rate of Deposit of the Drops on the Surface and the Percentage of Drops Caught. R. & M. 2025, 1940.